QDA Miner

Qualitative Data Analysis Software

User's Guide

Provalis Research

DISCLAIMER

This software and the disk on which it is contained are licensed to you, for your own use. This is copyrighted software owned by Provalis Research. By purchasing this software, you are not obtaining title to the software or any copyright rights. You may not sublicense, rent, lease, convey, modify, translate, convert to another programming language, decompile, or disassemble the software for any purpose. You may make as many copies of this software as you need for backup purposes. You may use this software on more than one computer, provided there is no chance that it will be used simultaneously on more than one computer. If you need to use the software on more than one computer simultaneously, please contact us for information about site licenses.

WARRANTY

The QDA Miner product is licensed "as is" without any warranty of merchantability or fitness for a particular purpose, performance, or otherwise. All warranties are expressly disclaimed. By using the QDA Miner product, you agree that neither Provalis Research nor anyone else who has been involved in the creation, production, or delivery of this software shall be liable to you or any third party for any use of (or inability to use) or performance of this product or for any indirect, consequential, or incidental damages whatsoever, whether based on contract, tort, or otherwise even if we are notified of such possibility in advance. (Some states do not allow the exclusion or limitation of incidental or consequential damages, so the foregoing limitation may not apply to you). In no event shall Provalis Research's liability for any damages ever exceed the price paid for the license to use the software, regardless of the form of claim. This agreement shall be governed by the laws of the province of Quebec (Canada) and shall inure to the benefit of Provalis Research and any successors, administrators, heirs, and assigns. Any action or proceeding brought by either party against the other arising out of or related to this agreement shall be brought only in a PROVINCIAL or FEDERAL COURT of competent jurisdiction located in Montréal, Québec. The parties hereby consent to in personam jurisdiction of said courts.

COPYRIGHT

TRADEMARK

TABLE OF CONTENTS

Introduction to QDA Miner 2.0

QDA Miner is an easy-to-use qualitative data analysis software package for coding textual data, annotating, retrieving and reviewing coded data and documents. The program can manage complex projects involving large numbers of documents combined with numerical and categorical information. QDA Miner also provides a wide range of exploratory tools to identify patterns in codings and relationships between assigned codes and other numerical or categorical properties. Documents are stored in Rich-Text Format and support font and paragraph formatting, graphics and tables. Documents may be edited at any time without affecting the existing coding.

QDA Miner can import and export documents, data and results in numerous file formats (MS Word, WordPefect, RTF, PDF, HTML, XML, MS Access, Excel, Paradox, dBase, etc.). It also provides unique integration with advanced quantitative content analysis, text mining (WordStat) and statistical analysis (Simstat) tools, providing easy combination and integration of qualitative and quantitative methods.

Installation Instructions

To install QDA Miner on your computer from the installation CD, place the CD in your CD drive. The installation program should start automatically. If it does not start, double-click on the "My Computer" desktop shortcut and then on the CD drive icon to display the content of the CD. Double-click on the INSTALL.EXE icon on the CD. Just follow the setup direction on screen.

If you downloaded the software, simply double-click on the SetupQM.EXE icon and follow instruction displayed on your screen.

Updating the Software

Minor revisions of QDA Miner are likely to be released on a regular basis and made available on the Provalis Research Web site. These releases may include bug fixes as well as minor improvements and sometimes even new features. Unless otherwise indicated, if your version of QDA Miner has the same starting number as the version available on the Web, you are entitled to a free update. For example, if you purchased the version 1.0 of QDA Miner, and the current version available on the Web site is 1.31, you should be able to update to this version for free.

To update your software to the latest version, select the WEB UPDATE command from the HELP menu or go to the QDA Miner Web site and select the QDA Miner WEB PAGE from the HELP menu, or go to the following web page:

www.provalisresearch.com/QDAMiner/QDAMiner.html.

From there, go to the download section of the site, download the trial version of QDA Miner and install this version over your existing one. QDA Miner will automatically register the trial version and transform it into a fully functional version.

Project Structure

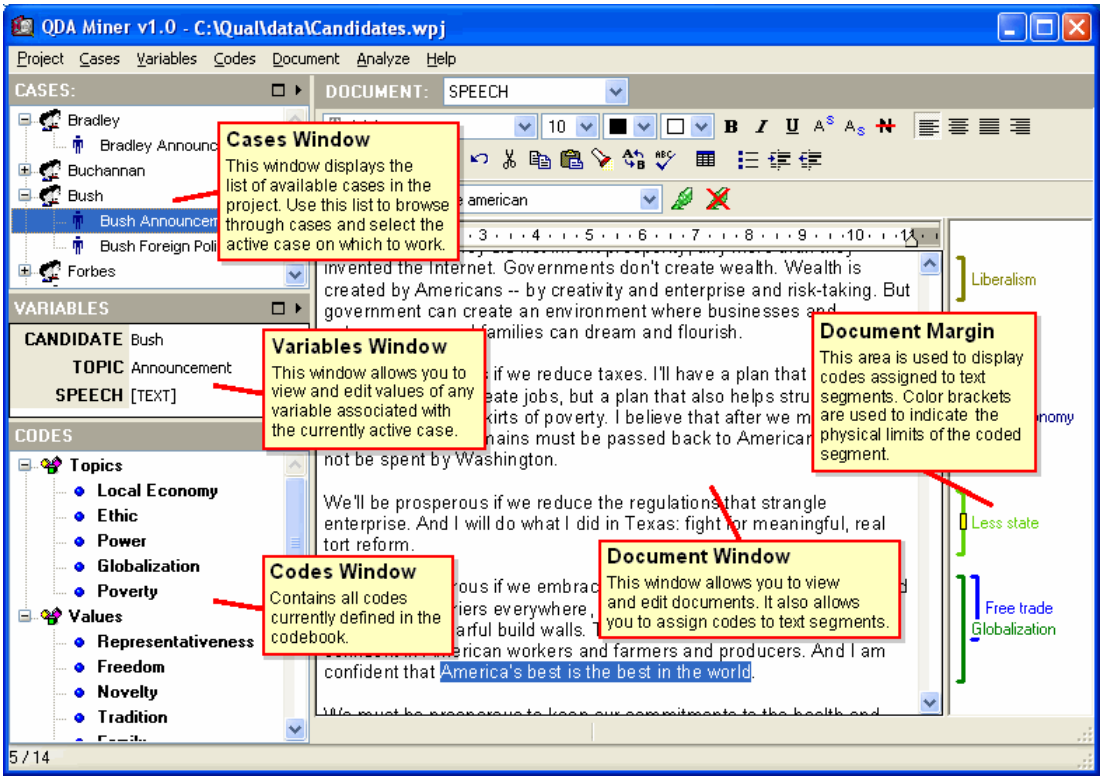
QDA Miner keeps all documents, coding schemes, codes, and notes in a set of files called a “project”. QDA Miner projects consist of multiple cases. A case is the basic unit of analysis of a project. It typically represents an individual, an organization, or a group. A case can contain several documents as well as numerous alphanumeric, numeric, categorical, date or Boolean variables. These variables are used to specify the properties associated with a case.

For example, if you want to analyze transcripts of in-depth interviews of numerous individuals, you may end up creating a project file where each case contains information associated with a single interviewee. One or several document variables may be created to contain transcripts of these interviews. You may also add other variables to specify sociodemographic information for the interviewee, group membership, etc, as well as the interview date and individual responses to closed-ended questions. Up to 2035 variables can be associated with each case. The number of cases in a single project is limited by the disk space, up to a maximum of 4 to 8 gigabytes.

One of the unique features of QDA Miner is its ability to explore relationships between any one of these properties and codes manually assigned to documents. For example, you can assess how the content of an interview is related to the interviewee’s gender or age, or how it relates to specific answers to a close-ended question.

The Working Environment

The QDA Miner working environment provides numerous tools to manage, view, edit, and code documents. It also gives access to a variety of retrieval and analysis tools. The picture below provides an example of the QDA Miner workspace.



The QDA Miner desktop consists of a menu bar at the top and four windows. The first three windows are located on the left side of the screen. These screens are, from top to bottom:

- The CASES window, which contains the list of available cases in the project. These cases are displayed either as a list or grouped as a tree. This window is used mainly to browse through cases and select the active case on which you want to work.

- The VARIABLES window, which displays and allows editing of values for any variable associated with the currently active case.

- The CODES window, located in the lower left corner of the screen, which contains all codes currently defined in the codebook.

The large DOCUMENT window on the right side of the screen is the main working space. This window is used to view and edit documents and assign codes to text segments. When a case contains more than one document, you can switch between document by selecting the document name from the list box located in the top left corner of this window. To the right of this window is a gutter area. This area is used to display codes assigned to text segments. Color brackets are used to indicate the physical limits of the coded segment. This area is used to review coding, remove or change assigned codes, and attach notes to any coded segments (see Working with Code Marks).

Each window may be resized either horizontally or vertically by dragging its border. The three windows in the right panel may be maximized to take up most of the available vertical space by clicking on the ☐ button located in the top right corner of the window. Clicking on the 🗗 button restores the maximized window to its original size. You can also minimize a single window by clicking on the ▶ button. The "folded" window may then be restored by clicking on the ▼ button.

The Main Menu

The PROJECT drop-down menu gives you access to various operations related to entire projects, such as project creation, opening and configuration. This menu also includes commands related to the import, export and backup of project files.

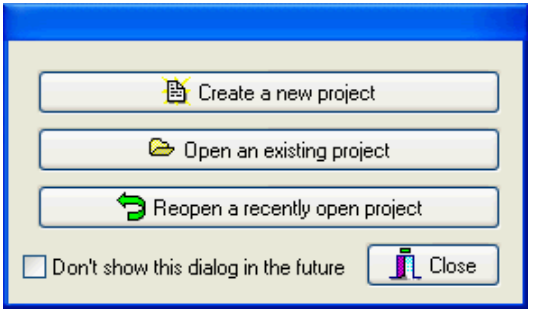
Operations associated with the main four windows of the working environment are grouped under menus with a corresponding name. For example, all operations relevant to the CASES windows are grouped under the CASES drop-down menu, while all commands relevant to the document currently displayed are located under the DOCUMENT drop-down menu.

The text and code retrieval features as well as the various exploratory tools available in QDA Miner (coding sequences or co-occurrences analysis, heatmaps and correspondence analysis plot, etc.) are all grouped under the ANALYSIS drop-down menu. This menu also provides access to external modules such as WordStat (quantitative content analysis and text mining) and Simstat (statistical analysis).

Opening an Existing Project

Opening a Project from the Introductory Screen

By default, QDA Miner stores data files in a folder named Projects. The Projects folder is located in the QDA Miner program folder. When you start QDA Miner, an introductory dialog box appears:



QDA Miner keeps track of the ten most recently use project files.

To open a recently opened project:

- Click on the **Reopen a recently open project** button and select its name from the displayed drop-down list.

To open another project:

- Click on the **Open an existing project** button. The **File** dialog box will appear.
- Locate and select your project (a QDA Miner project is stored with a .WPJ file extension).
- Click on the **Open** button.

Opening a Project When the QDA Miner Program Is Already Running

To open an existing project:

- Select the OPEN command from the PROJECT menu. The **File** dialog box will appear.
- Locate and select your project (a QDA Miner project is stored with a .WPJ file extension).
- Click on the **Open** button.

To open a recently opened project:

- Select the REOPEN command from the PROJECT menu
- Click on the project file that you want to reopen.

Opening a Project from Windows Explorer

You can also open a project by double-clicking its icon in Windows Explorer:

Creating a New Project

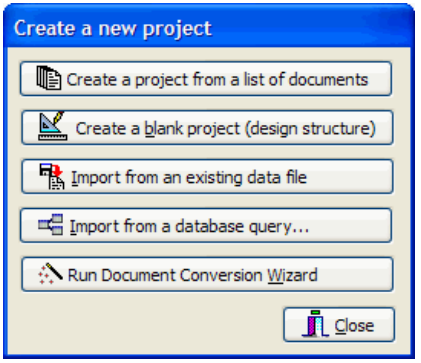
There are four ways of creating a QDA Miner project: (1) You can create a simple project by importing one or several documents; (2) you can create a new project from scratch by creating a project structure and then manually entering data and documents; (3) you can import existing data and documents stored in another file format such as Excel, MS Access, dBase, Paradox, etc.; or (4) you can also use the Document Conversion Wizard utility program to import several documents at once, perform some transformation and data extraction, and store them in a new project.

Creating an Project from a List of Documents

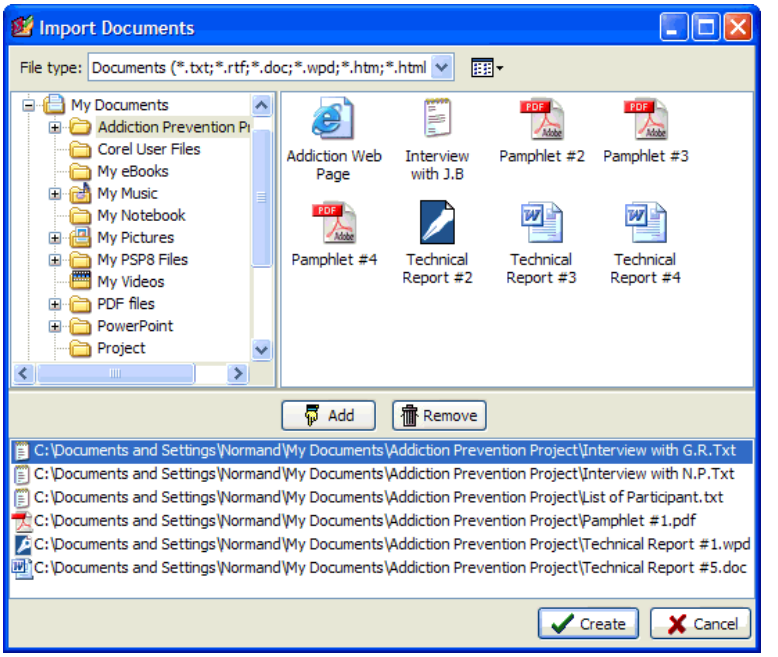
The easiest method to create a new project and start doing analysis in QDA Miner is by specifying a list of existing documents and importing them into a new project. Using this method creates a simple project with two variables: A DOCUMENT variable containing the imported documents, and a categorical variable containing the original name of the files from which the documents originated. Imported documents are stored in different cases so, if 10 documents have been imported, the file will have 10 cases with two variables each. To split long documents into several ones or extract numerical, categorical, or textual information from those documents and store them into additional variables, use the Document Conversion Wizard.

To create a new project using this method

- Select the NEW command from the PROJECT menu. This command calls up a dialog box similar to the one below.



- Click on the **Create a project from a list of documents** button. A dialog box similar to the one below will appear:



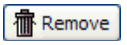
- Click on a folder in the folders list on the upper left section of the dialog box to display its contents. If you want to see the contents of a drive, go to the folders list, click on My Computer, and then double-click on a drive.

- In the upper right section of the dialog box, QDA Miner displays all supported document file formats that may be imported, such as MS Word, WordPerfect, RTF, PDF documents, plain text files or HTML. To display only documents of a specific type, set the **File Type** list box to the desired file format.

- Click on the file you would like to import. To select multiple files, hold down the **CTRL** key while clicking on the other files.

- Click on the 

button to add those documents to the list of documents to import, located at the bottom of this dialog box. You may also drag the files from the top right section to this list.

- To remove a file from the list of documents to import, select that file name and click on the 

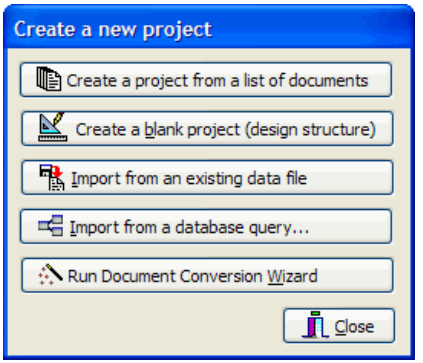
button.

- Once all files have been selected, click on the **Create** button. You will be asked to specify the name of the project that you want to create. If a project with an identical name already exists, you will be asked to confirm that you wish to overwrite the previous version of the project.

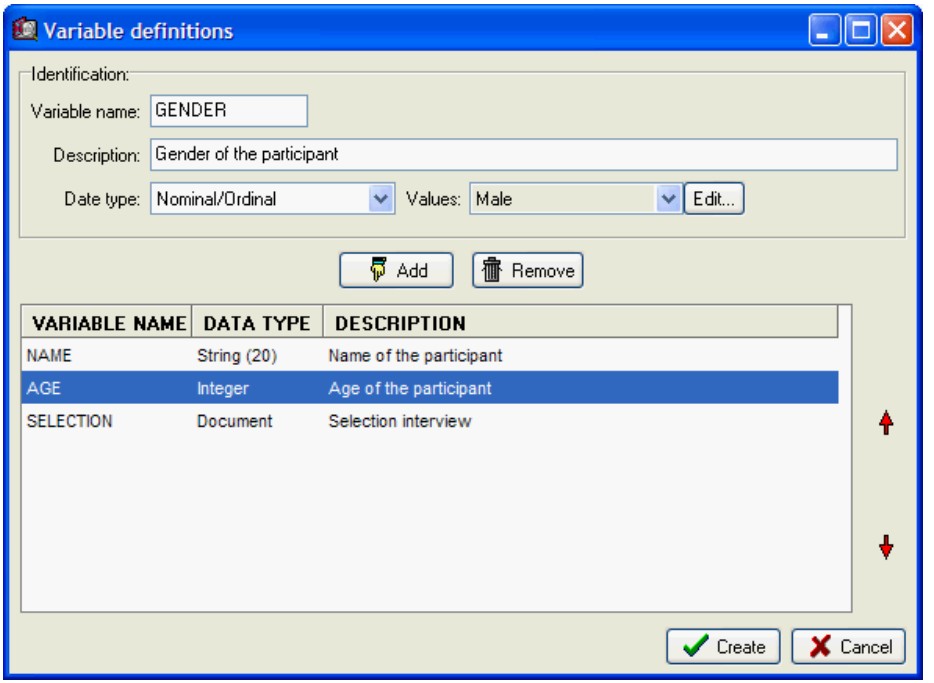
QDA Miner closes this dialog box, imports all the specified files into a new project and then brings you back to the main window.

Creating a Blank Project

To create a new data file, select the NEW command from the PROJECT menu. This command calls up a dialog box similar to the one below:



Select the **Create a Blank Project** button. The **Variable Definitions** dialog box will appear.



The first step involved in creating a new project file is to define the initial structure of the project file. This structure is defined by the list of variables that each case will contain. A variable may contain a document to be manually coded, but can also consist of a numeric value, a date, a Boolean value (true or false), etc. A single project can contain up to 2035 variables per case. It is possible to create several document variables for each case. The ability to store many documents per case is especially useful when the project involves a

set of several document types. A document variable may also be created to store notes or comments that are specific to a case. QDA Miner can later be instructed to search for and analyze specific document variables or in all documents.

In the **Variable Definitions** dialog box (see above), you can define various attributes for these new variables, such as their name, and whether they will contain numeric, alphanumeric values, or dates.

VARIABLE NAME - The first edit box at the top of the dialog box allows you to enter a variable name. Each variable name must be unique (within that project file). Valid variable names begin with a letter and may contain letters, numbers or underscore characters. Punctuation marks, blank spaces, accentuated and other special characters are not permitted. The maximum variable name length is ten characters.

DESCRIPTION - The Description option is used to enter a variable label that describes the content of the variable in more detail. You may leave this column blank if you wish since it is always possible to add or edit a description later using the VARIABLES | PROPERTIES command.

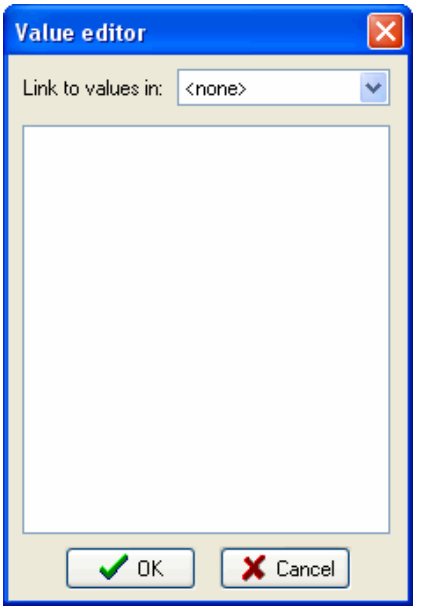
DATA TYPE - Each variable in the data file must have a type. QDA Miner supports the following types:

Document - This data type is used to store documents that will be manually coded. QDA Miner stores text in this data type using Rich Text Format (RTF). This format enabled the use of different fonts and styles and paragraph formatting. Graphics and tables may also be inserted in the document. Numerous file formats may be directly imported into document variables, such as plain ASCII files (*.TXT), Rich Text files (*.RTF), MS Word documents (*.DOC), HTML (*.HTM or *.HTML), Adobe Acrobat (*.PDF), and WordPerfect documents (*.WPD) (see **Importing a Document in a Case**, on page 40).

Numeric - Numeric variables can contain either integer or floating-point numbers. When you choose floating point numbers, you will be asked to specify the number of decimal places to display. Floating-point numbers are stored in the data file using double precision values (at least 15 significant digits). The decimals option is used exclusively to control how numeric values are displayed in the **Variables** grid and in other locations and does not affect the internal precision of the variable.

Nominal/Ordinal - Nominal or ordinal variables are used to hold a limited number of short strings used to describe specific properties of a case. For example, you may choose to use this variable type to identify the gender of the interviewee ("male" or "female") or its group membership ("manager", "employee", "client", etc.). You may also use this data type to hold ordinal ranking ("novice", "intermediate", or "expert") or responses to close-ended questions such as Likert scale items (e.g.: from "strongly disagree" to "strongly agree"). You should use this data type instead of short strings if you plan to analyze this variable or examine the relationship between its values and any other variable or coding of documents.

When a nominal/ordinal data type is selected, you will be asked to first provide a list of values that this variable can take. To enter new values, click on the **Edit** button. A dialog box similar to the one below will appear:



You can start typing values (one value per line) in the large edit box. If the current variable uses the same values as another existing variable, you may also establish a link to this other variable so that they will share the same list of values. Click **OK** to confirm the setting of these values. Note: This list may later be modified to add new values or edit existing ones.

Date - The date type holds a year, month and day. The display and data entry format used for dates is based on the Windows date setting.

Boolean - The Boolean type stores a value that can be either true or false (or Yes or No).

Short String - Short string variables can contain up to 254 alphanumeric characters. When creating a short string variable, you must specify first the maximum number of characters this variable with hold.

To create a new variable:

- Enter a unique variable name.
- Enter a description (optional).
- Select the data type for this variable.
- Set the option associated with the chosen data type (see above).
- Click on the **Add** button to add the defined variable to the list.

To remove a variable from the list:

- Select the row containing the variable you want to delete.
- Click on the **Remove** button.

To change the position of a variable in the list

- Select the row containing the variable you want to move.
- Click on the **up** or **down** arrows located to the right of the grid until the variable appears in the desired location.

To create the data file:

When you have finished defining the structure of the new project, click on the **OK** button.
You will be asked to specify the name of the project that you want to create. If a project with an identical name already exists, you will be asked to confirm the overwriting of the older project.

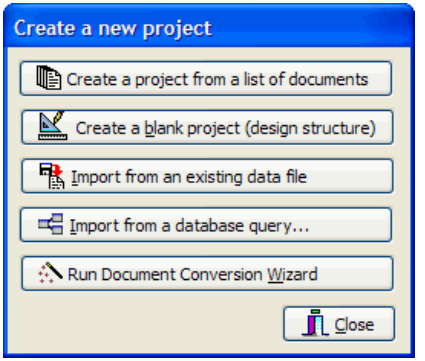
Importing an Existing Data File

QDA Miner allows you to directly import data files from spreadsheet and database applications, as well as from plain ASCII data files (comma or tab delimited text). The program can read data stored in the following file formats:

- MS Access
- DBase
- Paradox (v3.0 - v5.0)
- Lotus 123 (v1.0 - v5.0)
- MS Excel
- Quattro Pro (v1.0 - v6.0)
- Comma Separated Values
- Tab Separated Values
- Triple-S XML file (interchange standard for survey data)

To import data from any of these applications:

- Select the NEW command from the PROJECT menu. The following dialog box will appear:



- Click on the **Import from an existing database** button.
- Select the file format using the **List File of Type** drop-down list.
- Select the file you want to import and click **OK**.

We will now discuss the specific file formats in greater detail, including formatting requirements (if applicable), supported features and limitations.

ASCII Data Files

QDA Miner will read up to 500 numeric and alphanumeric variables from a plain ASCII file (text file). The file must have the following format:

- Every line must end with a carriage return.
- The first line must include the variable names, separated by spaces, tabs and/or commas.

- Variable names may not be longer than ten characters. Longer strings are truncated at ten characters.
- The remaining lines must include numeric scores separated by spaces, tabs and/or commas.
- Each line must contain data for one case; variables must be in the same order for all cases.
- All invalid scores and all blanks encountered between commas or tabs are treated as missing values. A single dot can also be used to represent a missing value.
- Comments can be inserted anywhere in the file by putting an asterisk ('*') at the beginning of the line.
- Blank lines can also be inserted anywhere in the file.

Spreadsheet Data Files

In spreadsheet programs, you can enter both numeric and alphanumeric data into the cells of a data grid. QDA Miner can import spreadsheet files produced by EXCEL, LOTUS 1-2-3 (v1.1 to v5.0) and QUATTRO PRO (v1.0 to v6.0).

When you select a spreadsheet file format for importing, the program displays a dialog box in which you can specify the spreadsheet page and the range of cells where the data are located. You must specify a valid range name or provide upper left and lower right cells, separated by two periods (such as A1..H20). If you set the **Range Name** list box to **ALL**, the program attempts to read the whole page.

Formatting spreadsheet data

The selected range must be formatted sothat the columns of the spreadsheet represent variables (or fields) while the rows represent cases or records. As well, the first row should preferably contain the variable names, while the remaining rows should hold the data, one case per row. QDA Miner will automatically determine the most appropriate format based on the data that it finds in the worksheet columns. Cells in the first row of the selected range are treated as field names. If no variable name is encountered, QDA Miner will automatically provide one for each column in the defined range.

Database Files

dBase and Paradox files

QDA Miner can import dBase and Paradox data files. However, the length of alphanumeric fields should not exceed 256 characters and memo fields are not supported. If you do have this type of data, you may use the exporting capabilities of your database program to create a data file that is more compatible with QDA Miner (such as a FoxPro 2.x data file or a tab delimited text file).

MS Access data files

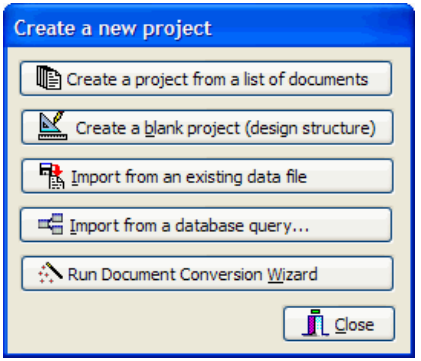
When you select a MS Access data file, the program displays a dialog box in which you can specify the table where the data are located. Once the table has been selected, click on the **IMPORT** button to import the data file.

Importing Existing Documents as a New Project

The Document Conversion Wizard is a utility program used to import one or more documents into a new project file. This tool supports the importating of numerous file formats including plain ASCII, Rich Text Format, MS Word, HTML, Acrobat PDF files, and WordPerfect documents. It may be instructed to split large files into several cases and to extract numeric and alphanumeric data from these files. The Document Conversion Wizard may be run either as a standalone application or from within QDA Miner. When no splitting, transformation or extraction of information from documents are necessary, one can also use the easier to use **Creating a Project from a List of Documents** method.

To Run the Document Conversion Wizard from Qda Miner:

- Select the NEW command from the PROJECT menu. The following dialog box will appear:



- Click on the **Run Document Conversion Wizard** button.

The program then guides you through the necessary steps needed to import the documents and extract relevant data. For more detailed information on the document importing process using this utility program consult the **Document Conversion Wizard** help file.

Project Description and Notes

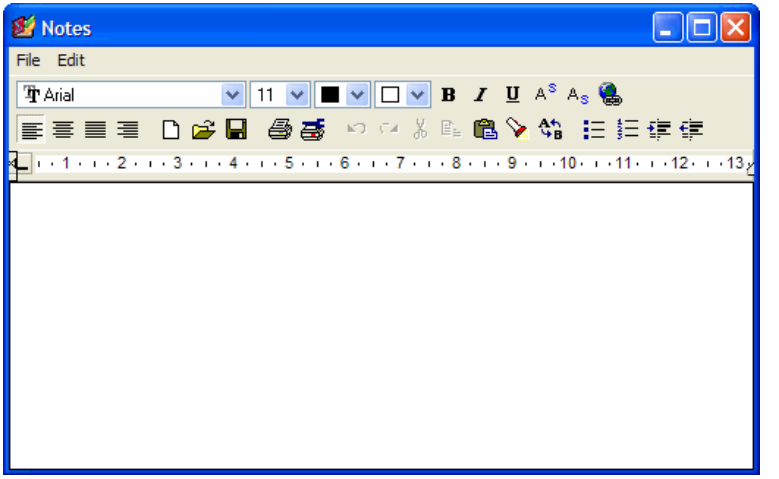
QDA Miner allows you to specify a description for the project as well as write general notes to yourself or to be shared with other people working on this project. The project description typically contains general information about the project and may be configured so that it will be displayed automatically when opening the project file. Notes may be used as reminders of decisions you made, things to be done, etc.

To enter a project description:

- Select the PROPERTIES command from the PROJECT menu.
- Enter a description for the project.
- To display this description automatically when the project file is opened, select the **Show Description upon Opening** checkbox.
- Click on the **OK** button to save the changes you made.
- To quit this dialog box without saving the changes, click on the **Cancel** button.

To enter or edit general notes:

- Select the NOTES command from the PROJECT menu, or press the F3 key from anywhere in the program. This will open a Note editor like the one below:



- Enter your comments or observations. You can format text and paragraphs, change font attributes, insert graphics and tables. You may also save your notes to disk in RTF format.

- To close this dialog box, select the CLOSE button from the PROJECT menu.

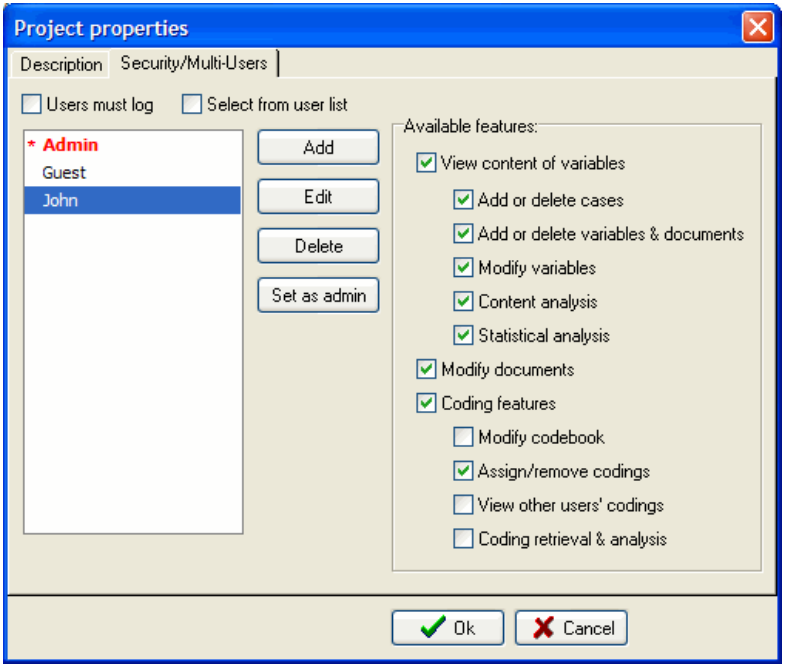
Setting Security & Multi-users Settings

It is possible to limit the number of people who can access and edit a project or limit the type of operations that can be performed by certain individuals by creating user accounts and requiring people to provide a user name and password when they access the project. This feature is useful for preventing deletion of existing cases or variables, or the editing of documents or values stored in other variables. The multi-user account feature is also useful for assessing the reliability of coding by allowing the same documents to be coded by different users (see **Assessing Inter-coders Agreement** on page 120).

When setting a project to support multiple users, one of these users should be able to control access to the project, create and delete user accounts and define passwords. This user is commonly known as the administrator. When creating a new project, an administrator account is automatically created. Both the default user name and password is ADMIN. If you choose to restrict access to some users, it is highly recommended that you change this user name and password to prevent unauthorized changes to user access rights.

To change the multi-users settings of a project:

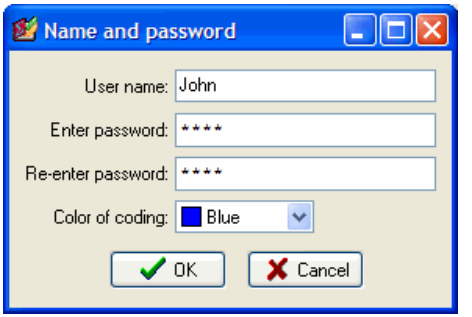
- Select the PROPERTIES command from the PROJECT menu.
- Click on the SECURITY / MULTI-USERS tab. The following dialog box will appear:



By default, opening a project without using the user log screen gives the user all administrator access rights. Enabling the **Users must log** option will display a logon dialog box prompting the user to enter a valid user name and password in order to access the project file. The user name can be either entered in a text box or selected from a drop-down list box. To display existing user names in a list box, set the **Select from users list** option.

To add a new user account:

- Click on the **Add** button The following dialog box will appear:



- Enter in the user name.
- In the **Enter Password** edit box, enter this user's password.
- Enter the password a second time in the next edit box to make sure it was entered correctly.
- QDA Miner can emphasise coding assigned by different coders using a distinct color for each coder for displaying their code marks. The Color of Coding option can be used to select the color used to display codes assigned by this coder (see **Modifying Code Marks Color Scheme** on page 59).
- Click **OK** to save the new user account.

Once a user's account has been created, you have to specify which specific features this user will be able to access.

To define the user access rights:

- Select the user for which you want to define or edit access rights.
- In the list of available features to the right of the dialog box, select the features you want this user to have access to and clear the features that you do not want him to access.

When creating several user accounts to establish inter-coders agreement, it is strongly recommended that you deactivate the **View other users' coding** option. This prevents a user from seeing code marks assigned to text segments by other users. Disabling this feature also restricts any coding retrieval procedures to the coding created by this user and prevents access to the inter-coders agreement dialog box. To prevent users from modifying existing documents or values stored in variables, clear the box beside the **Modify variables** and **Modify documents** options. It may also be a good idea to disable the **Add or delete cases** and **Add or delete variables/documents** options. Lastly, if you want to prevent users from adding new codes to the codebook, or deleting or modifying existing ones, make sure the **Modify Codebook** option is disabled.

To remove a user account:

- In the list of existing accounts to the left of the dialog box, select the account you want to remove.
- Click on the **Delete** button.
- Select **Yes** to confirm the deletion of this user's account.

To save the changes you made to the accounts and close this dialog box, click on the **OK** button. To close this dialog box without saving any changes, click on the **Cancel** button.

Adding a New Variable

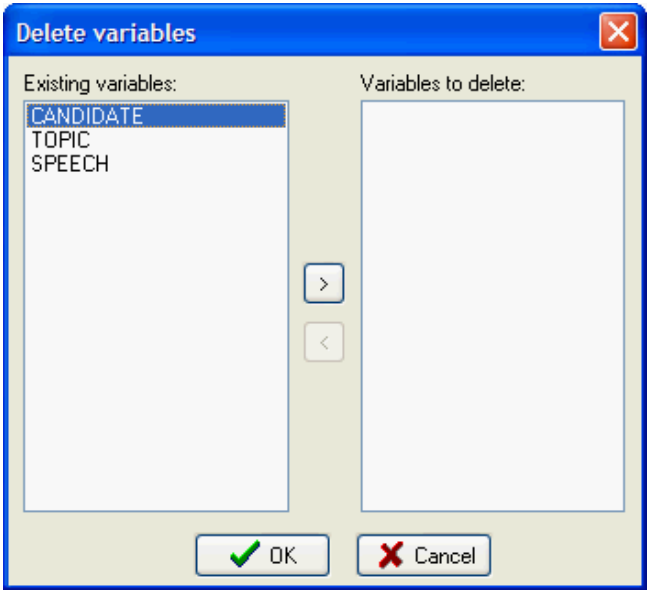
The ADD command in the VARIABLES menu is used to add new variables to the project file. This command displays a dialog box similar to the one used to create variables for a new data file (for more information on this dialog box and its options see **Creating a new empty project** on page 13).

Specify the name of the new variables and set the variable types, sizes and descriptions.

Click on the **OK** button to create these new variables and add them to the end of the current data set.

Deleting Existing Variables

To delete one or more variables, select the DELETE command from the VARIABLES menu. The following dialog box will appear.



Highlight the names of the variables that you want to delete and click on the  button to move them to the **Variables To Delete** list box.

To delete successive variables, click on the first variable, drag the mouse cursor down the list to highlight multiple variables, and then click on the  button.

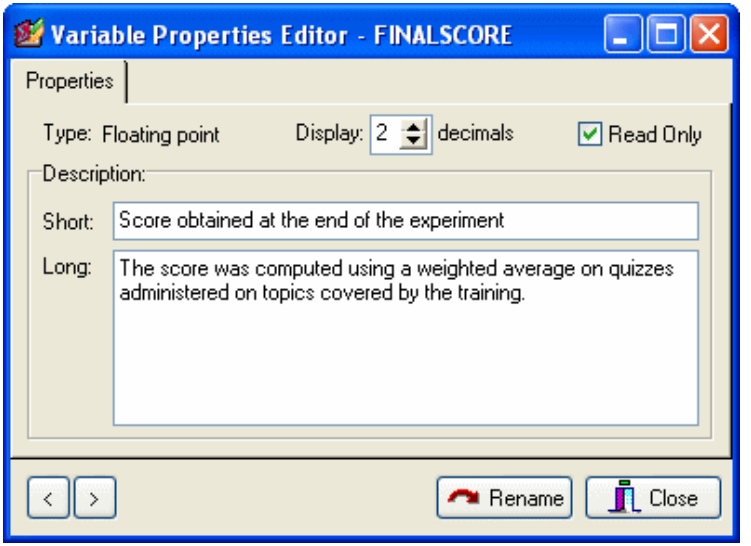
To remove a variable from thelist of variables to delete select the name of this variable from the **Variables to Delete** list box and click on the  button.

Click on the **OK** button to delete all selected variables.

Editing Variable Properties

QDA Miner allows you to edit various properties of existing variables in your project. For example, you can attach a short and a long description to a variable, set the number of decimal places for floating-point numerical values, edit values of categorical variables, etc. You can also set an individual variable as "read only" or change its name. To access the variable properties dialog box, use the following steps:

In the Variables windows, position the cursor on the variable that you want to edit.
Choose the PROPERTIES command from the VARIABLES menu, or right-click anywhere in the Variables window and select the EDIT PROPERTIES command. This displays the **Variable Properties Editor** dialog box as shown below.



Once in the dialog box, you can navitage through variables by clicking either the < or the > button. When viewing or editing the properties of a categorical variable, a second page appears. You can add new values, as well as edit or delete existing ones on this page (see below). The first page of the dialog box offers the following options:

READ ONLY - When selected, this option prevents a variable from being modified. This option is useful to prevent accidental or unauthorized changes to the values of a variable. To prevent the modification of data for an entire project file, see **Security & Multi-Users Settings** on page 21. Setting a variable to **read only** only affects the editing of values in existing cases, but still allows users to create new cases and assign values to these new cases.

DECIMALS - When the variable is a floating-point number, the decimals option is used to specify how many decimal places to display in the data windows. Floating-point numbers are stored in the data file using double precision values (at least 15 significant digits). The decimals option is used

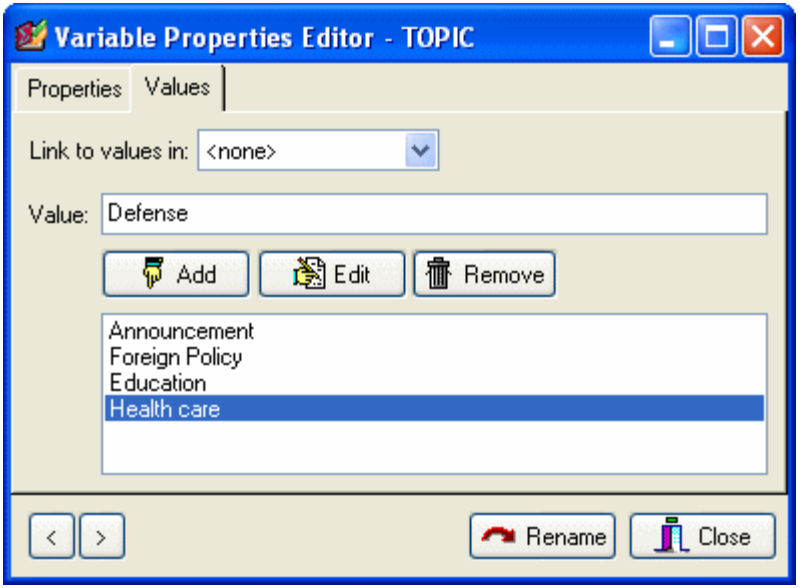
exclusively to control how numeric values are displayed in the **Variables** grid and in other locations and does not affect the internal precision of the variable.

DESCRIPTION - This option lets you enter both a short single-line alphanumeric description as well as a detailed description of the variable. The short description is displayed in various locations in the program to help remind users of the exact content of this variable.

RENAME BUTTON - You can use this button to change the name of the current variable. When you click this button, you will be prompted for a new variable name. This new name must not exist in the current data file and should follow the basic rules for valid variable names.

The Values Page

When viewing the properties of a categorical variable, a second page is shown. This **Values** page allows you to add new values, as well as edit or delete existing ones.



To add a new value labels:

- In the Value edit box, enter the string that will be used to describe this new value.
- Click on the **Add** button.

To remove an existing value label:

- In the list box located in the lower half of this page, select the label you want to delete.
- Click on the **Remove** button

To edit an existing value label:

- In the list box located in the lower half of the page, select the label you want to edit.
- Click on the **Edit** button.
- Once you finished editing the label, click on the **OK** button to apply the change.

Using an Existing Value Labels Definition

In some projects, several variables share the same value labels. For example, a questionnaire may use a common ordinal scale for several questions. Rather than re-entering the same value labels over and over again, QDA Miner can establish a link between a variable without value labels and an existing one that already contains labels.

To establish a link, click on the down arrow button of the **Link to values in** list box, then select from the list of variables the one containing the value labels you want to use. These labels will appear in the **Value** list box.

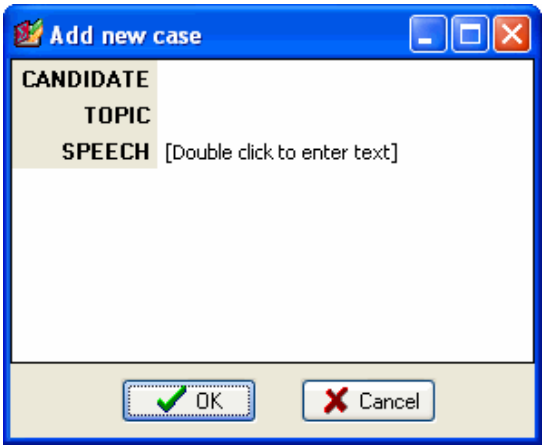
Once a link has been established, every change made to the value labels list will affect the labels associated with the original variable as well as with all other variables currently linked to this variable.

You may also use this feature to copy value labels from one variable to another. To do this, follow the previous instructions to link the current variable to the one from which you want to copy labels. The labels should appear in the value labels list. Then remove this link by setting the **Link to values in** option to none. You may now edit the newly copied labels without affecting the labels of other variables.

Adding and Deleting Cases

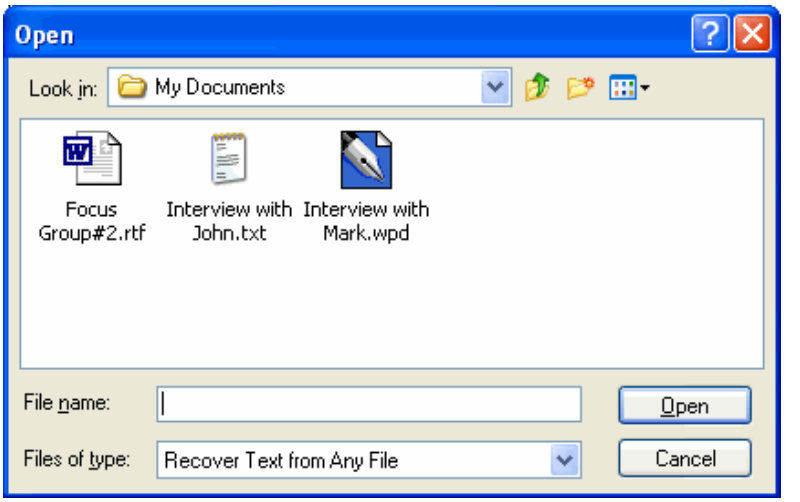
Adding a New Case

To create a new empty case, select the ADD command from the CASES menu. A data entry form will appear allowing you to enter values for each variable in the current project.



To enter a new value, click on the date entry cell located to the right of the variable name that you want to edit. If the variable is numeric or alphanumeric, you can start typing the data you want to store in this variable. For categorical variables, dates and Boolean values, press the **F2** key or double-click on the cell to display the list of available values or open a date editor. The tab key will take you to the next variable.

To enter text or import an existing file into a document variable, double-click on the data entry cell or press **F2**. A text editor window will appear. You can start entering the text you want to store in this document variable. You can also import an existing document by selecting the OPEN command from the FILE menu or by clicking on the button. When you select this command, an **Open File** dialog box will appear.



Select the type of file you want to import. QDA Miner can read the following document types:

RTF files (Rich Text Format)
HTML documents
ANSI files (or plain ASCII files)
MS Word documents
Windows Write documents
WordPerfect documents
Acrobat PDF files

Other file formats such as InWriter/Notetaker and PocketWord may also be supported, depending on your system installation.

Once you have finished editing the document, select the CLOSE command from the PROJECT menu or click on the  button located on the upper right corner of the editor.

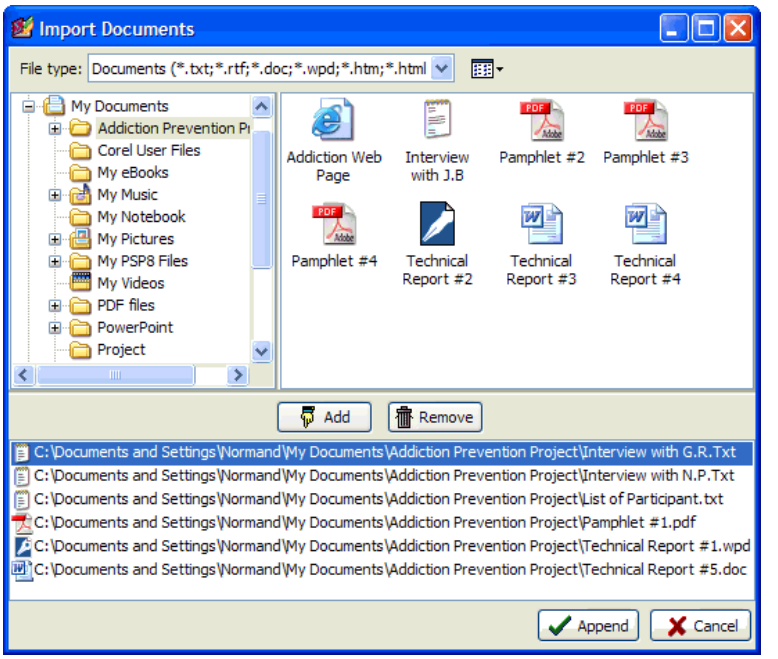
Click on the **OK** button to confirm the creation of this new case. To exit this dialog box without creating this new case, simply click on the **Cancel** button.

Deleting a Case

To delete a case, simply move to the case you want to delete by selecting its descriptor in the CASES window and then select the DELETE command from the CASES menu.

Appending New Documents

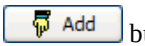
To append documents and store them in new cases, select the APPEND DOCUMENTS command from the CASES menu. A dialog box similar to the one below will appear:



- Click on a folder in the folders list on the upper left section of the dialog box to display its contents. If you want to see the contents of a drive, go to the folders list, click on **My Computer**, and then double-click on a drive.

- In the upper right section of the dialog box, QDA Miner displays all supported document file formats that may be imported, such as MS Word, WordPerfect, RTF, PDF documents, plain text files or HTML. To display only documents of a specific type, set the **File Type** list box to the desired file format.

- Click on the file you would like to import. To select multiple files, hold down the **CTRL** key while clicking on the other files.

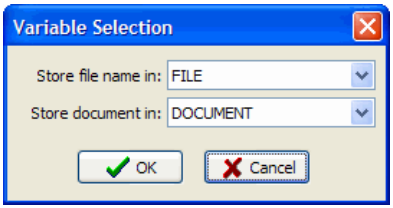
- Click on the 

button to add those documents to the list of documents to import, located at the bottom of this dialog box. You may also drag the files from the top right section to this list.

- To remove a file from the list of documents to import, select that file name and click on the 

button.

- Once all files have been selected, click on the **Append** button. If the project contains more than one categorical or document variable, a dialog box similar to this one will appear:



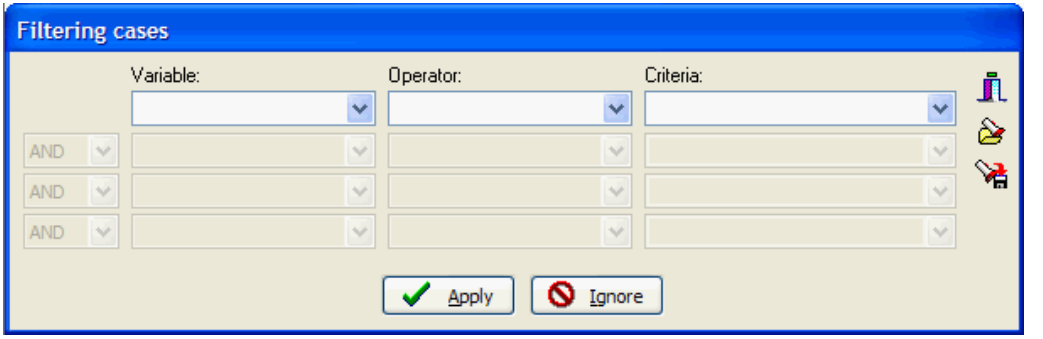
- Select the categorical variable in which the file name will be stored. Set this list box to <none> to prevent the program from storing this information in the project.

- Select the document variable where the imported documents should be stored.

- Click on the **OK** button.

Filtering Cases

The FILTER command in the CASES menu temporarily selects cases according to some logical condition. You can use this command to restrict analysis to a subsample of cases or to temporarily exclude some subjects. The filtering



condition may consist of a simple expression, or include up to four expressions joined by logical operators (i.e., AND, OR).
The following table shows the various operators available for each data type

DATA TYPE	AVAILABLE OPERATORS
NOMINAL / ORDINAL	Equals Does not equal Is empty Is not empty
NUMERIC and DATE	Equals Does not equal Is greater than Is lesser than Is greater than or equal to Is lesser than or equal to Is empty Is not empty
BOOLEAN	Is true Is false
STRING	Contains Does not contain Is empty Is not empty
DOCUMENT	Is empty Is not empty

Once a filtering expression has been entered, you can apply the filter and leave this dialog box by clicking on the **Apply** button. If the filter expression is invalid, a message will appear exiting from the dialog box will not occur.

To temporarily deactivate the current filter expression, click on the **Ignore** button. The filter expression will be kept in memory and may be reactivated by selecting the FILTER CASES command again and clicking **Apply**.

To store the filtering expression, click on the  button and specify the name under which those filtering options will be saved.

To retrieve a previously saved filter, click on the  button and select from the displayed list the name of the filter you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

To exit from the dialog box and restore the previously active filtering expression, click on the **Close** button.

Exporting Selected Cases

The PROJECT | EXPORT command saves a copy of the current project under a different name or exports the project to another file format. QDA Miner supports the following export formats:

- DBase
- Paradox
- Lotus 123
- Excel
- Quattro Pro
- Comma Separated Values
- Tab Separated Values
- Triple-S XML (interchange standard for survey data)
- XML

When exporting a data file, QDA Miner will use the current filtering condition to determine which cases will be exported. The active sorting order will also be used to control the case sequence in the new file.

To export data to any of these applications:

- Set the filtering (see page 32) and sorting conditions (see page 36) of the active data file to display the cases as they should be exported.
- Select the EXPORT | PROJECT FILE command from the PROJECT menu.
- Select the file format you want to create using the **Save As Type** drop-down list.
- Enter a valid filename with the proper file extension.
- Click on the **Save** button.

Creating a File with Subsets of Cases

The export feature can also be used to create a new data file with a structure identical to the current data file, but with only a subset of cases.

To create a project with a subsets of cases:

- Set the filtering and sorting conditions of the active data file to display the cases as they should be saved in the new data file.
- Select the EXPORT | PROJECT FILE command from the PROJECT menu.
- Set the **Save As Type** drop-down list to QDA Miner Project
- Enter a valid filename.
- Click on the **Save** button.

Exporting Documents

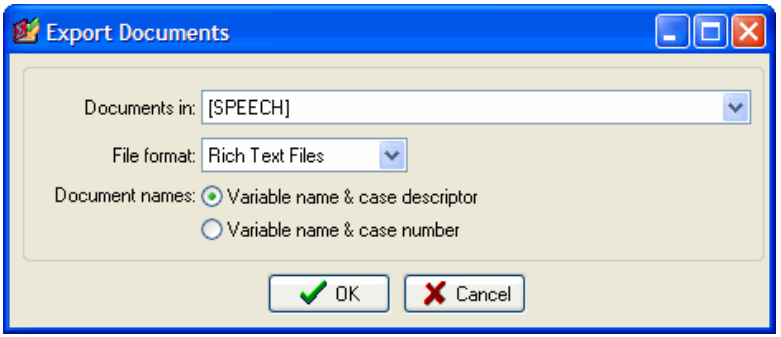
The EXPORT | DOCUMENTS command saves all documents in a project file associated with one or several document variables into separate files. QDA Miner supports the following document formats:

- Rich Text
- HTML
- Plain text file

When exporting documents, QDA Miner will use the current filtering condition to determine which documents will be exported.

To export documents to disk:

- Set the filtering condition of the active data file to display the cases containing the documents you want to export. If you want to export all documents, remove any filtering condition.
- Select the EXPORT | DOCUMENTS command sequence from the PROJECT menu. The following dialog box will appear:



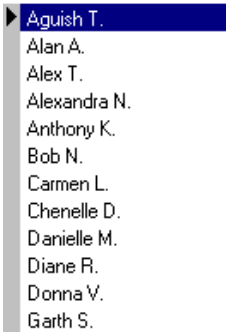
- Select the variables containing the documents you would like to export to disk.
- Set the **File Format** option to the desired format.
- In the **Document Names** option, select the method that should used to automatically create file names.
- Click on the **OK** button.

  You will be asked to specify a folder under which document files should be stored.

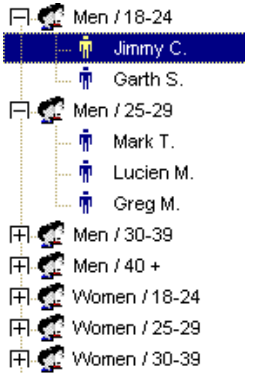
Setting the Cases Descriptor and Grouping

The CASES window shows all currently active cases in the project. By default, cases are listed in their order of creation and are identified by their physical position in the data file (e.g. case #1, case #2, etc.). However, you can edit the list order, group cases in categories, and define a custom descriptor that will be used to identify a case based on the values contained in one or more variables.

There are two major ways to display the active cases. These cases may be displayed either in a single sorted list or grouped by category based on the values of one or two variables. The example below shows a list of cases sorted alphabetically on interviewee name.



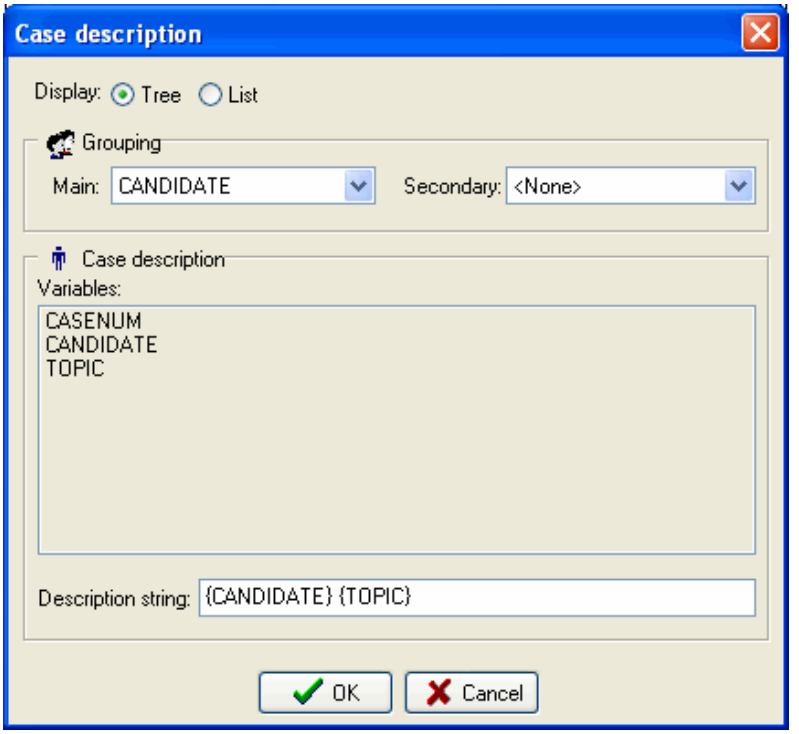
The second example below shows the same cases grouped by gender and age.



In this last example, all cases appear in a tree in which each node represents the combined value of the grouping variables. You can display or hide cases under this category by clicking on the + or - sign located to the left of the node or by double-clicking on the node itself. To display or edit a specific case, simply select its case descriptor.

TIP: Displaying cases grouped in a tree means that the program must read all cases in the project, which may cause some delays when working with large data files containing several thousand cases. For this reason, when working with such a large project, it may be preferable to display cases as a list.

To adjust the descriptors used to identify cases or to choose how these cases will be organized, select the GROUPING/DESCRIPTOR command from the CASES menu. The following dialog box will appear:



DISPLAY - This option lets you select whether cases will be displayed in **TREE** mode or as a single **LIST**.

When you choose to display cases in a tree, you will be asked to provide up to two variables that will be used to group cases. The **MAIN** list box allows you to select the first grouping variable, while the **SECONDARY** list box allows you to further break down the cases into the subcategories defined by these two variables. For example, if you choose GENDER as the main grouping variable and AGE as the second grouping variable, and if we assume that the AGE variable contains only three values (16, 17 and 18), cases will be grouped under the following six categories:

MALE - 16
MALE - 17
MALE - 18
FEMALE - 16
FEMALE - 17
FEMALE - 18

If you choose to display cases in a single list, cases will be displayed by default in order of creating. The MAIN and SECONDARY list boxes may however be used to re-order these cases and display them sorted in ascending order on one or two variables.

CASE DESCRIPTION - This section of the dialog box allows you to specify a label that will be used to describe each case. The label may be changed by editing the text in the **DESCRIPTION STRING** edit box. To insert the value stored in a specific variable into the description, simply enter the variable name in uppercase letters and enclose this name between braces. Alternatively, you can insert a variable name at the current caret location by clicking on the corresponding item in the VARIABLES list located just above the edit box.

If you enter the following string:

```
{GENDER} subject - {AGE} years old
```

The {GENDER} and {AGE} strings will be replaced with their corresponding value for this specific case. If the current case contains information about a seventeen-year-old male, the above string will be displayed as:

Male subject - 17 years old

Beside the name of variables, it is also possible to insert the following string:

```
{CASENUM}
```

This string will display a unique case number, representing the physical order of this case in the project file.

Importing a Document into a Case

The Document window allows users to enter text directly in the document editor or paste text from the clipboard. It is also possible to import an existing document stored on disk into the project.

To import a document into an existing case:

- From the **Cases** window, select the case in which you would like to store the document.
- If your project contains more than one document variable per case, make sure the Document window points to the proper document by selecting its name from the list box on the title bar of this window.
- Select the DOCUMENT FILE | IMPORT command from the DOCUMENT menu or click on the  button on the document editor toolbar. An **Open File** dialog box will appear.
- Select the file format of the document that you want to import from the **Files of Type** list box located at the bottom of this dialog box.
- Highlight the file name that you want to import and click **Open**.

Please note that importing a document will overwrite existing text stored in the currently selected document. To append a file to an existing text, simply select this text, copy it to the clipboard, import the new document, and paste into the text back in the proper location.

To import a document into a new case:

- Select the ADD command from the CASES drop-down menu. A data entry dialog box will appear.
- Double-click on the document variable to access the text editor.
- Select the OPEN command from the PROJECT menu or click on the  button on the editor toolbar. An **Open File** dialog box will appear.
- Select the file format of the document you that want to import from the **Files of Type** list box located at the bottom of this dialog box.
- Highlight the file name that you want to import and click OPEN.
- Exit the editor by selecting the CLOSE command from the FILE menu or clicking on the  button.
- After entering data for other variables, click on the **OK** button.

Archiving a Project

During a research project, data files are frequently modified (cleaning, coding, etc.). At some point you may need to return to a previous version of an existing data file to recover lost variables or cases that has been transformed or deleted, or just to make routine verifications. In order to do this, several successive backup copies of this data file should be created and kept. Another reason to make backup copies of data files is to prevent the accidental loss of an entire data file caused by a hardware failure or a software malfunction.

QDA Miner provides a simple archiving procedure that allows users to quickly create backup copies. This procedure stores a copy of the currently active project and all its related files (structure definition, documents, codebook, value labels, etc.) in a single compressed file. QDA Miner uses the industry standard ZIP format as its own archive file format so that you can easily manage these archives files outside QDA Miner using any application that can manipulate these files. As well, the high level of compression achieved on typical QDA Miner data files (usually more than 90-95%) allows you to create backup copies of large projects on a single diskette, or keep several copies of your data file on your hard drive without sacrificing precious disk space.

Another benefit of this archiving feature is to facilitate the transfer of your data file and all its related files to another computer by creating a single file that includes all associated files.

To make an archived copy of the current data file:

- Select the MAINTENANCE | BACKUP | CREATE command sequence from the PROJECT menu. An **Export Data** dialog box will appear.
- Set the drive and directory setting to the location inwhich you want to store the compressed file.
- Enter a valid file name and click **OK**.

To restore a file from an archived copy:

- Select the MAINTENANCE | BACKUP | RESTORE command sequence from the PROJECT menu. An **Open File** dialog box will appear, displaying all ZIP files.
- Select the proper ZIP file and click on the **OK** button. A second dialog box will appear, prompting you to specify the location in which data should be restored. The default location is the same directory as the archive file.
- If needed, change the drive and directory setting and click on the **OK** button to proceed with the extraction.

Merging Projects

The QDA Miner MERGE command allows one to combine two or more project files into one. This operation is achieved by opening a "master" project that will act as the recipient and then selecting external project files from which information will be retrieved. The MERGE command will perform up to 4 types of data import in a single operation.

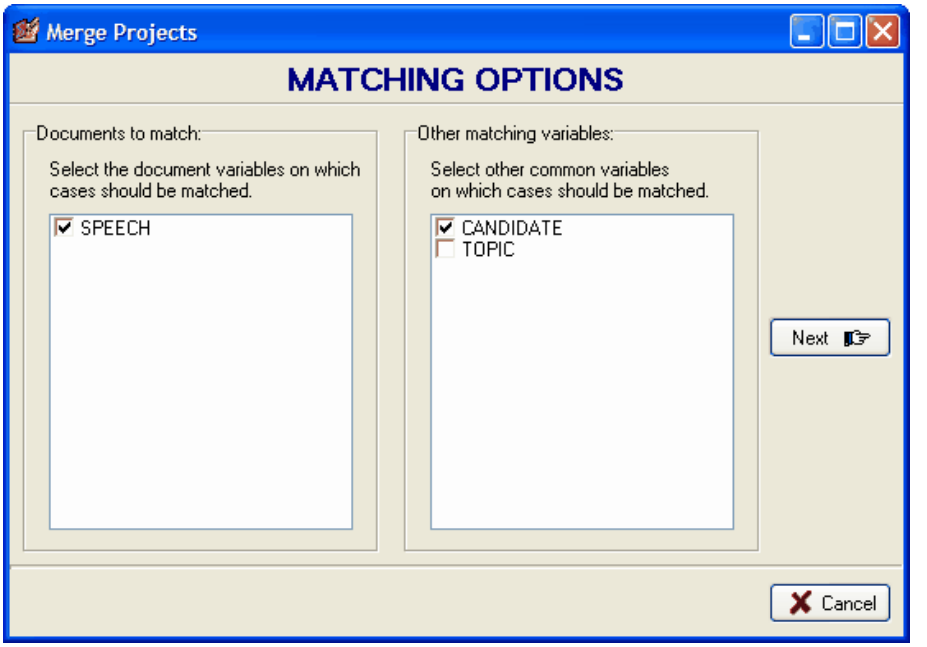
- It will attempt to match cases and documents and to import for common documents any code assignment that is not already in the master project.
- If necessary, it will also import the codebook from the external file and merge its codes and categories into the existing codebook (see also **Importing a Codebook** on page 53 for the single importation of codebooks from other projects).
- If the program cannot match cases in the imported project to existing ones in the main project, it can optionally append those cases along with their associated documents and codings.
- If the imported project contains unique variables, the program will allow selecting from those variables the ones that should be appended to the existing master project.

This merging feature is useful to allow one to synchronize one's own work performed on different computers. As well, both an individual or various team members may work on distinct sets of documents and then merge them along with their associated codings into a single master project.

The Merge feature may also be used by a team of researchers working on the same set of documents. They can perform their own coding and editing on their own copy of a master project file, working on different computers at different times. Once individual coding and data editing has been done, they can then combine all codes assigned by each researcher into a single project file. Such a scenario may be used to separate the amount of work among several people. It is also useful to assess the reliability of coding made by different coders, by allowing different researchers to code independently the same sets of documents (for more information how to assess the reliability of coding, see **Assessing Inter-Raters Agreement** on page 120).

To merge two project files:

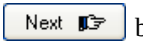
- Open the project in which the information currently contained in other project files will be stored.
- Select the MAINTENANCE | MERGE command from the PROJECT menu. An Open File dialog box will appear.
- Select the second project from which to import codes, coding, cases, or variables. The following dialog box will appear:



In order to merge coding made to duplicate documents stored in different projects files, QDA Miner needs to decide whether two documents are the same. This is performed by comparing the first 200 alphanumeric characters in each document (if the document is shorter, its entire text is compared). Spacing, parentheses and punctuation marks, as well as any other formatting characters including tabs and end-of-lines, are ignored. While it is unlikely that two documents will start exactly the same way, such a situation may still occur, especially when dealing with very short responses. To prevent such a situation from occurring, it is possible to instruct the program to consider a match when two cases share identical values on any other variables found in both project files.

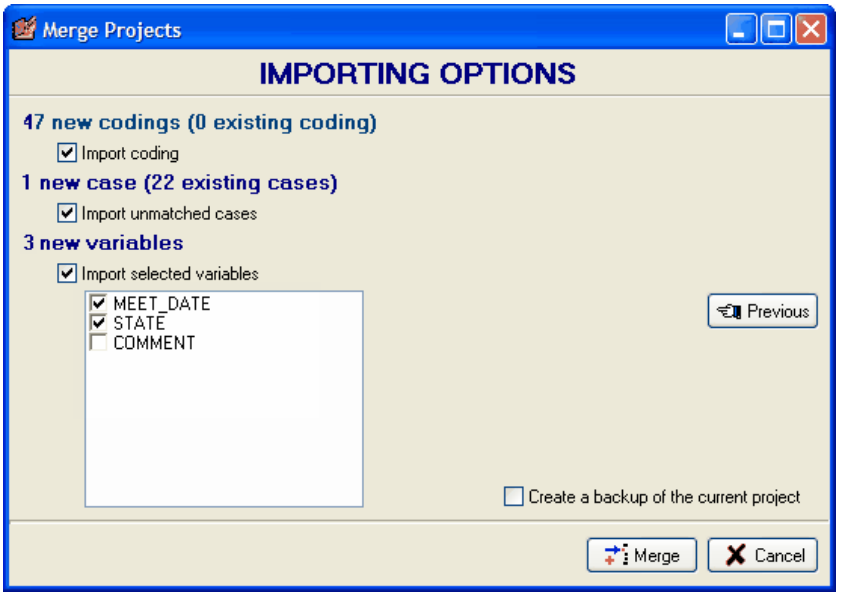
The first list box on the left side of the dialog box is used to specify on which documents this pairing should be made. Since coding may only be imported for documents that have been previously matched, this list box also specifies where the codings to be imported are located.

After selecting at least one document variable, select from the list of all variables that are common to both files any other variable on which the pairing of cases should be made. Two cases will be considered identical only if all the selected documents and all the values of those selected variables are the same.

Once the matching options are set, click on the 

button. This will instruct QDA Miner to apply the matching criteria, check for possible conflicts, and then move to the second page of the dialog box. If two records in the master project share the same values on all selected variables, QDA Miner will display an error message stating that the master project contains duplicate entries. If the selected key variables are unique to each case, then the program will move on to the next page.

The second page of the dialog box previews the predicted outcome of the merging operation and specifies what should be imported. Depending on the nature of the information stored in the second project file, up to three import options will be available:



Import coding - While attempting to match cases on common documents, QDA Miner also compares all the codings applied to those documents. If it finds codings that do not exist in the master project, it will offer the option to import those codings. The number of new codings that would result from a merge is reported at the top of the second page of the dialog box, as is the number of common codings found in both project files. Select the IMPORT CODING box to instruct QDA Miner to import those additional codings. Note that this option only applies to documents contained in both project files. Any new document imported by the addition of new cases or new variables will always be merged into the current project along with all its associated codings.

Import unmatched cases - After attempting to pair all cases in both project files, QDA Miner will report the number of cases in the second project for which no match has been found. Select this option to append those cases at the end of the current project.

Import selected variables - If some variables are unique to the imported project file, it will be possible to integrate all or some of those unique variables into the main project by enabling this option and selecting from the list of variables located below those that should be imported.

Merging several projects into a master file may result in several important changes to the original project file. Some of those changes may not correspond to what was expected. Since it could be very difficult to reverse all the changes resulting from a merge and recover the original project file as it was before this operation, it is highty recommended either to perform this operation on a copy of the original project file or to create a backup of your original project before merging any data into it. While the creation of backups may be performed at any time in QDA Miner by using the MAINTENANCE | BACKUP command, an

option has been added to this dialog box to facilitate the access to the BACKUP command. Enabling this option will instruct the program to first perform a backup of the current project file before importing begins.

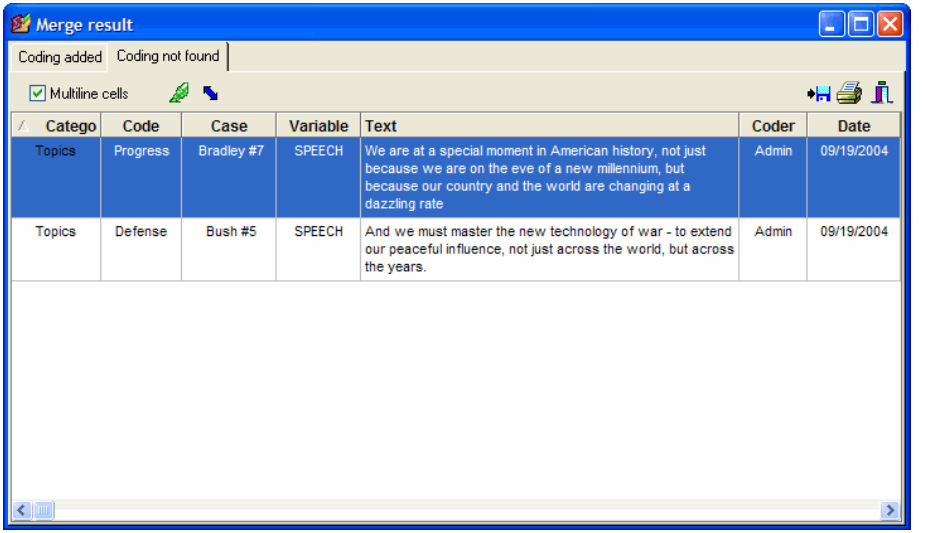
To cancel the merging operation and return to the QDA Miner main screen, click on the Cancel button.

To go back to the first page of the dialog box to adjust matching settings, click on the Previous button.

Once ready, click on the Merge button to begin merging the projects. If the backup option has been set, you will be presented with a dialog box that will ask you to specify the name of the backup file. The program will then proceed to the import of all the information and will display a message summarizing the changes made to the master project.

Reviewing imported coding

When QDA Miner imports codings, it attempts to match the original text and location of a coding in the imported project to a segment of the master document that points to the exact same text and the same or nearby position. However, it may sometimes fail to find the proper location, usually because one of the two documents has been modified. After importing codings, QDA Miner displays a dialog box with a list of all codings added to existing documents as well as any coding that could not be assigned to the documents in the master project (see: "Coding not found" tab below).

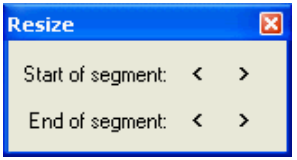


The first tab of this dialog box displays all the codings that have been added to the master project. The second tab displays any coding which QDA Miner was unable to locate. This second table of unmatched coding includes the text to which this code had been originally assigned in order to make the match manually.

Selecting an item in this table, either by clicking it or by using the keyboard cursor keys (such as the UP or DOWN arrow keys), will automatically display the corresponding case and document name in the main window and will select a portion of the document text located near the code. If slight changes have been made to either version of the document, the highlighted segment in the main window should be relatively close to the correct location.

To manually adjust the location of an unmatched code:

- Click on the  button. The selection resizing tool will appear:



- From this dialog box, click on the corresponding arrow button to move the starting or ending location of the selected text in the main window. Clicking a **<** button moves the chosen limit of the selection to the left and up while clicking a **>** button moves this limit to the right and down. Repeat the same move several times by clicking and holding the mouse button down. Release the mouse button when done.

- Once the selection has been adjusted to closely match the text originally associated with this code, click on the  button to assign this code to the new selection.

Clearing all code assignments

A researcher may want to start over with a fresh uncoded project or create a copy of the original project without any coding in order to allow someone else to code the same set of documents. To clear all codings made to the current project:

- Select the MAINTENANCE | CLEAR ALL CODINGS command from the PROJECT menu, and click **Yes** to confirm the deletion of all code assignments.

- As an extra precaution, a dialog box will be displayed requiring the typing of the word **YES** in an edit box. Typing **YES** and clicking **OK** will result in the removal of all codings and their associated memos.

To remove specific codings, see **Removing Coded Segments** on page 56.

Retrieving Misplaced Codes

When a code is assigned to a selected segment of a document, the code itself is stored within this document at the exact location where it was assigned. The associated coding will always be preserved, even if a portion of text or an entire document is copied to the clipboard and pasted somewhere else (whether in the original project or another one) or if it is saved to a disk as an RTF document. This method of storing codings instead of keeping an index of their location has the advantage of allowing one to modify the original document without the risk of corrupting the coding index. However, one inconvenience is the possible insertion within a project of documents already containing codings for which there is no corresponding code in the current project codebook. A software or hardware malfunction causing an abrupt termination of the program may also result in the loss of some codes in the codebook or in some codings being left without a corresponding code.

The RETRIEVE MISPLACED CODES command instructs QDA Miner to browse through all the documents in the current project and search for any coding without an associated code in the codebook and to append any new code to the existing codebook. All new codes will be stored under a special category named **Recovered**.

To run this command, select the MAINTENANCE | RETRIEVE MISPLACED CODES command from the PROJECT menu. The program immediately browses through all documents and reports the number of codes that have been retrieved.

If the retrieved codes are obsolete or no longer valid, simply erase the code in the codebook, which will result in the removal of all codings associated with it (see **Deleting Existing Codes** on page 50).

Managing the Codebook

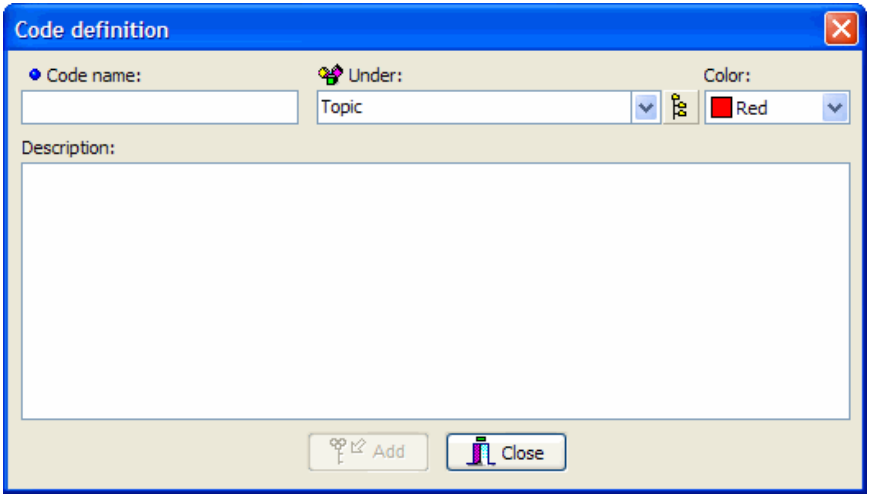
The main function of QDA Miner is to assign codes to selected text segments and then analyze these codes. Available codes associated with a project are located in the CODES window in the lower left corner of the application workspace. Codes are grouped into categories. The list of available codes, known as the codebook, is displayed as a tree structure, in which categories are the nodes under which associated codes can be found.

You can expand a category to view all associated codes or close it by clicking on the plus or minus sign located to the left of the category name. When a new project is created, this window is empty.

This section provides basic instructions on how to add new codes, organize them into categories and edit and delete existing ones. It also provides instructions on how to change the location of a code in the codebook, how to merge a code into another and how to import an entire codebook from another project.

Adding a Code

To add a new code to the existing codebook, select the ADD command from the CODES menu. The following dialog box will appear:



When you create a new code, you must (I) specify a unique name and (ii) select a category.

The **UNDER** list box allows you to select the category under which this code will be stored. This control can be used both as an edit box to create a new category and as a list box from which you can select an existing category.

If you want to add the new code to an existing category, select the category name from the list of available categories by clicking on the down arrow key to the right of the list box and by selecting the category name.

To add the new code under a new category, simply enter the name of the new category. By default, the new category is listed at the top level of the codebook. To create a subcategory underneath another one, type its full path separating each level by a backslash. For example, typing `Topics\Economy` will create a subcategory Economy under the Topics main category. If this top-level category does not exist, the program will create both this category and its subcategory and then add the specified code into this subcategory. To store a new subcategory into an existing path, one may also type this new category name and then click on the  button, which will display the tree view of the codebook categories. From that, select the parent category under which this new category will be stored. Selecting any existing category from this tree view automatically inserts the full path before the new category name in the edit box.

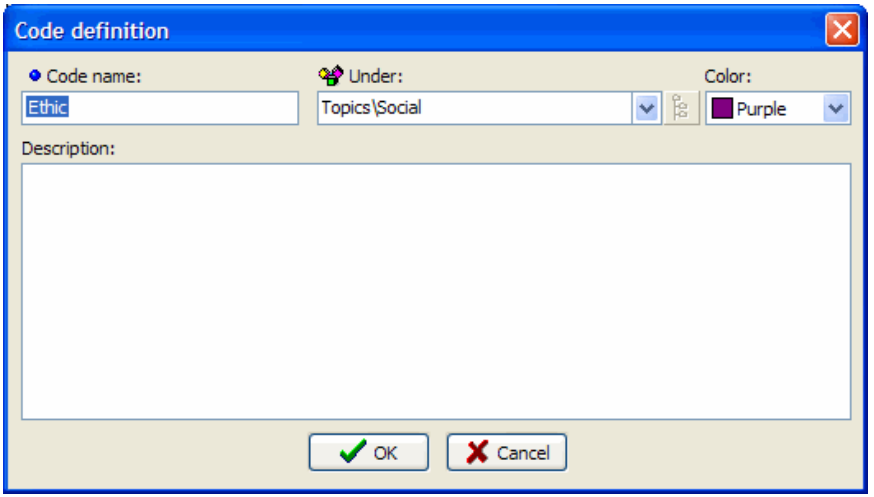
The maximum number of levels a codebook can contain is set to 4 by default for every new project. Since the last level can only contain codes, the actual maximum level of a subcategory is equal to this maximum minus 1. To increase or decrease the maximum number of levels allowed in a project codebook, run the PROPERTIES command from the PROJECT menu, and set the **Maximum Levels in the Codebook** option to a number between 2 and 8.

When a code is assigned to a text segment, a bracket appears in the MARGIN of the document to indicate the beginning and end of the code along with the code name. The **COLOR** option can be used to select the color of this bracket and of its associated label.

The **DESCRIPTION** option allows you to enter a definition or a detailed description of the code. You can use this section to specify coding instructions for the coders, along with examples or non-examples, and related codes that may be used in place of or in conjunction with this code.

Modifying Codes

To modify an existing code or category, simply select the code or the category that you want to edit and then select the EDIT command from the CODES menu. If the currently selected item is a category, you will be prompted to provide a new name for this category. If a code is selected, the CODE DEFINITION window will appear, allowing you to change the code name, its category, its description or the color used to identify segments that has been assigned to this code:



Once finished, you can click **OK** to apply the changes, or **Cancel** if you do not want to save any changes.

Changing the name of a code will trigger QDA Miner into browsing all documents in the active project in order to apply this change.

Deleting Existing Codes

To delete an existing code, simply select the codes you want to delete and then select the DELETE command from the CODES menu. You can also delete all codes associated with a category by selecting this category instead of a code. If you attempt to delete a category, you will be prompted to confirm the deletion of all codes within this category.

Deleting one or more codes from the codebook results in the removal of the corresponding codes from all documents associated with the current project. If you want to delete an existing code but assign another code to all segments currently associated with this code, use the code merging feature (see below).

Moving Codes

Codes and categories in the codebook may be freely moved in the codebook using one of the following methods:

To move a code using drag & drop:

- Click on the code or category that you want to move, holding down your mouse button.
- Drag the item to its new location and release the mouse button. How the dragged item is inserted in the target position depends on the exact position where you drop it.
    - If the item is dropped on a category name, it will be stored as the last item under this category.
    - Dropping a code over another code name inserts the dropped item before the target code.

To move a category and its items using drag & drop:

- Click on the category that you want to move, holding down your mouse button.
- Drag the category to its new location and release the mouse button. If the item is dropped on another category name, it will be inserted before the target category. If it is dropped over a code, the item will become a subcategory of this code category.

  QDA Miner will not allow a category to be dropped in a location underneath it. It will also forbid dropping a category at a location if the number of levels resulting from such a move exceeds the highest level allowed for this project codebook.

  To increase or decrease the maximum number of levels allowed in a project codebook, run the PROPERTIES command from the PROJECT menu, and set the **Maximum Levels in the Codebook** option to a number between 2 and 8.

Merging Codes

During the analysis process, you may choose to merge several codes together. QDA Miner offers you the possibility of merging a single code into another or of merging all codes in a single category into one code. When merging a code into another one, QDA Miner removes the first code from the codebook and automatically recodes all text segments tagged with the first code by giving them a new code. For example, if you decide to merge the code FAMILY into TRADITION, QDA Miner will delete FAMILY from the codebook. It will then search for all documents in the project for text segments that have already been tagged with the FAMILY code and change the assigned code to TRADITION. When merging all codes of one category, the program removes all those codes, transforms the category into a new code, and then automatically recodes all text segments tagged with any of the deleted ones by this new code. If a segment has been tagged using two or more of the merged codes or when two of those coded segments overlap each other, the coded segments are automatically consolidated so that only one tag is attached to the text segments.

To merge a code into another one:

- In the CODES window, select the code you want to eliminate.
- Select the MERGE INTO command from the CODES menu. The following dialog box will appear:

- Select which code you want to attach to text segments already assigned to the code you have decided to delete.
- Click on the **OK** button to proceed.

Splitting a Code

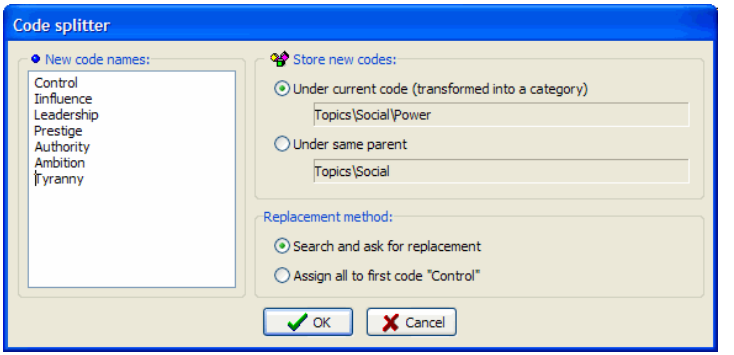
Sometimes, one may feel the need to refine a code that has been created at an early stage of analysis in order to make an important distinction one didn't initially consider or to introduce some nuance that would allow a finer-grained analysis. Frequency analysis of codings may also reveal that a code definition is too broad and occurs too often to be of any value for the interpretation of the data. The splitting code function of QDA Miner can be used to solve this. The procedure allows one to subcategorize a code into several other codes in a single step that involves:

1. The transformation of the initial code into a category.
2. The creation of the codes underneath this category
3. The assignment of all existing text segments associated with the initial code to either one of these new codes.

To split a code into several ones:

In the CODES window, select the code you want to split.

Select the SPLIT CODE command from the CODES menu. A dialog box similar to the one below will appear:



In the **New Code Names** edit box, enter the name of all the new codes into which the initial code could be recoded, one code name per line.

In the **Store New Codes** section of the dialog box, specify where those new codes will be stored. By default, the initial code is transformed into a category, and the new codes are stored under this newly created category. This can be achieved by selecting the **Under Current Code** radio button. One may instead wish to store all new codes in the same location as the initial code by selecting the **Under Same Parent** radio button. Please note that when the code to be split is at the deepest level allowed by the project setting, only this second option will be allowed. In that case, to make sure all new items will be grouped under a new category, one can either move the code to a higher level of the hierarchy prior to the execution of this split command (see **Moving Codes**, page 50), or increase the maximum number of levels allowed in this project codebook.

Next, select which replacement method will be used to replace all coded segments currently tagged with the initial code. Two options are available: (1) The program can search among all documents for any coded segment with the split code, display it and, for each coded segment that is found, ask the user to choose among all the new codes the one that should be attached to this segment; or (2) Choose to immediately replace all coded segments with the first code specified in the list of new codes. The user may return later to the coded segments associated with this first code in order to review those and, if needed, modify the code associated with those segments.

Click on **OK** to proceed to the splitting of the code.

If the **Search and Ask for Replacement** option was selected, the program will search among all documents and all cases for text segments that have been tagged with the initial code that has just been split. It will stop at the first occurrence, highlight the text segment, and display a dialog box similar to the one below.

In the Replace With list box, select the code that is appropriate to this text segment and click on the **Replace** button. The program will then search for the next occurrence and stop again, allowing one to select another code and assign it to this other text segment. The program proceeds this way until all text segments associated with the initial code have been assigned a new code. However, one may choose to stop this interactive search-and-replace procedure by choosing a code and clicking on the **Replace All** button, instructing the program to replace the initial code for all remaining text segments with a new code. The user may then return later to the coded segments associated with this code to review those and, if needed, modify the code associated with those segments, either individually by selecting their code marks and using the RECODE INTO command, or globally using the SEARCH & REPLACE command from the CODE menu (see **Searching and Replacing Codes**, page 59).

Importing a Codebook

Quite often codebooks developed for a specific project may be used again for other coding purposes. QDA Miner can import a codebook and merge its codes and categories into the existing codebook. When merging codebooks, codes or categories whose names already exist in the current codebook are simply ignored.

To import an existing codebook from another project:

- Select the IMPORT CODEBOOK command from the CODES menu. A **File** dialog box will appear.
- Select the project containing the codebook that you want to import and click on the **Open** button.
- You will be asked to confirm the merging of imported codes into the current codebook. Click **Yes** to confirm the merging of codes.

Coding Documents

The simplest and most obvious method of qualitatively code textual information is to read each document in a project and manually assign existing codes in a codebook to a selected segment of text. QDA Miner allows you to easily attach a code to a block of text that may be as small as a single character but that will more often consist of a few words, or one or more sentences or paragraphs. A code may even be assigned to an entire document.

The following section provides instructions on how to assign codes to text segments, resize the selected segment, change the associated code, attach a comment to a coded segments, as well as how to delete one or more codings.

Assigning Codes to Text Segments

To assign a single code to a segment of text:

You can use several methods to assign codes to text segments.

Method 1 - Quick assignments of codes to whole paragraphs

- Select the code from the code list (bottom left of the screen) by clicking it and holding the mouse button down.
- While holding the mouse button down, drag the mouse cursor over the paragraph you want to code.
- Release the mouse button to assign the dragged code to the paragraph found under the mouse cursor.

Method 2 - Highlight text and double-click on the code

- Highlight the text segment that you want to code.
- Double-click on the code in the code list.

Method 3 - Highlight text and drop the code

- Highlight the text segment you want to code.
- Drag an item from the code list to the selected segment.

Method 4 - Using the code list tool button.

- Select the proper code from the list box located in the toolbar above the document.
- Highlight the text that you want to code.
- Click on the  button.

The last used code can be quickly assigned to a newly selected text segment by right-clicking in the document and selecting the CODE AS command from the shortcut menu.

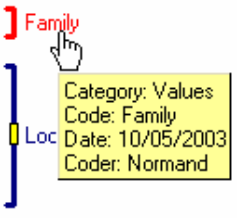
To Assign Several Codes to the Same Segment:

If you hold the **CTRL** key down while assigning a code to a text segment, the text segment will remain highlighted. You can then assign other codes to the same segment by double-clicking on them in the code list or by dragging them from the code list to the highlighted segment.

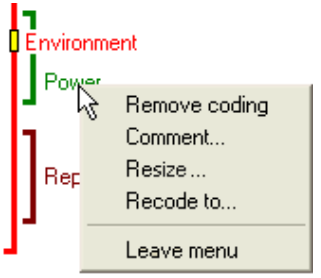
Working with Code Marks

When a text segment has been successfully coded, a code mark will appear next to it in the margin area on the right side of the document. Code marks looks like the following:

Color brackets are used to indicate the physical limits of coded segments. Moving the mouse cursor over a code mark (either the bracket or the code label) will display more detailed information about this specific coded segment, such as the category to which it belongs, its creation date, the name of the coder who created it, and any comment assigned to this coded segment.



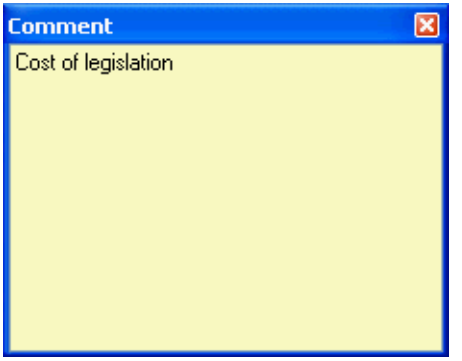
Clicking a code mark selects the text segment to which this code has been assigned. Clicking on the same code mark a second time opens a shortcut menu. This menu can also be accessed by moving over a code mark and right-clicking it.



Attaching a Comment to a Coded Segment

To attach a comment to a coded segment:

- Select the coded segment to which you want to attach a comment by clicking its code mark.
- Click a second time to display the shortcut menu.
- Select the COMMENT command. A small window like the one below will appear.



- Enter the text that you want to associate with the selected coded segment.
- Click on the **X** button in the upper right corner to save the comment and close the window.

When a note has been assigned to a coded segment, a small yellow square is displayed in the middle of the code bracket. To edit an existing code, follow the same steps. To remove a comment, simply open the note editor and delete all text in the note editing window.

To retrieve a list of all comments in a project file see page 135.

Removing Coded Segments

To remove coded segments:

Two methods may be used to remove codes assigned to text segments.

Method 1 - Using the coded segment panel

- Select the coded segment that you want to delete by clicking its code mark.
- Click this mark a second time or right-click to display a shortcut menu.
- Select the REMOVE CODING command.

Method 2 - Using the code deletion dialog box

The **Code deletion** dialog box allows you to delete one or more codes in the currently selected document.

- Click on the  button located on the toolbar of the document window. A list of all codes assigned to the current document will appear.
- Select the names of all the codes that you want to remove.
- Click on the **OK** button.

Note: The above operations are only used to remove code assignments to text segments. They do not result in the removal of the code from the codebook. To remove a code along with all associated assignments, see **Deleting Existing Codes** (on page 50). To remove all code assignments in a project see **Clearing All Code Assignments** (page 46).

Changing the Code Associated with a Segment

To change the code associated with a text segment:

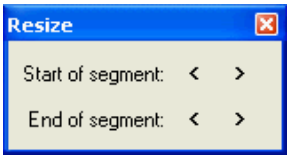
- Select the coded segment that you want to remove by clicking its code mark.
- Click a second time to display the shortcut menu.
- Select the RECODE TO command. A window containing all available codes will be displayed.

To change the code currently assigned to a new one, select the new code by clicking it, followed by the **OK** button. You can also double-click on the desired code.

Resizing Coded Segments

To resize a coded segment:

- Select the coded segment that you want to resize by clicking its code mark.
- Click a second time to display the shortcut menu.
- Select the RESIZE command. A small window like the one below will appear.



From this dialog box, you can move the starting or ending location of the selected segment by clicking on the corresponding arrow button. Clicking a < button moves the chosen limit of the segment back to the left and up while clicking a > button moves this limit forward to the right and down.

You can repeat the same move several times by clicking and holding the mouse button down. Release the mouse button when you are done.

Hiding code marks

Sometimes, one would rather focus on, or examine, specific coding or assigned codes. In such a situation, it may be preferable to hide unrelated code marks in the document margin.

To hide code marks associated with a specific code:

- In the codebook list, select the code for which the associated code marks will be hidden.
- Select the HIDE CODING command from the CODES menu.

Codes for which associated code marks have been hidden are displayed in the codebook using a lighter font and a grey bullet.

To hide code marks associated with a category of codes:

- In the codebook list, select the category containing all the codes for which the associated code marks will be hidden.
- Select the HIDE CODING command from the CODES menu.

To reveal hidden code marks:

- In the codebook list, select the code or the category for which the previously hidden code marks will be displayed.
- Select the HIDE CODING command from the CODES menu which will remove the check mark and will reveal all associated code marks.

Please note that hiding code marks only affects the display of those marks in the document margin. It has no effect whatsoever on the analysis or retrieval of coded segments.

Modifying Code Mark Color Scheme

The color scheme used to display code marks is set by default to highlight different codes. It may, however, be changed to emphasize the different coders who have assigned codes to the text segments.

To change the coding scheme to highlight different coders:

- Select the CODE MARKS COLOR command from the CODES menu.
- Choose the BY CODER menu item.

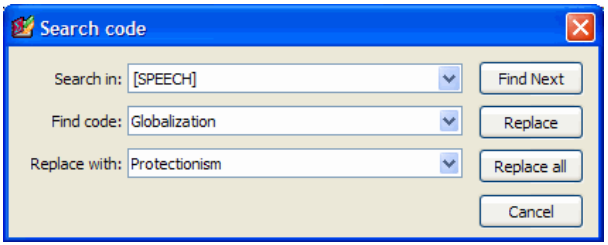
To change the coding scheme to highlight codes:

- Select the CODE MARKS COLOR command from the CODES menu.
- Choose the BY CODE menu item.

To modify colors associated with different codes, see **Modifying Codes** on page 49. For instructions on how to assign different colors to different coders, see **Security and Multi-Users Settings** on page 21.

Searching and Replacing Codes

The SEARCH & REPLACE command from the CODES menu may be used to interactively review codes assigned to text segments and, if needed, replace them with another code. When this command is executed a dialog box similar to the one below will appear:



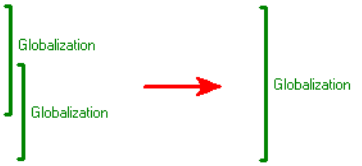
The **Search In** option allows you to specify which document variables to search. If the current project contains more than one document variable, you will have a choice of selecting either one or a combination of them. By default, all document variables are selected. To restrict the analysis to only a few of them, click on the down arrow key at the right of the list box. You will be presented with a list of all available document variables. Select the variables on which you want the search to be performed.

The **Find Code** option allows you to choose which code you want to find. In the **Replace With** list box, select the replacement code.

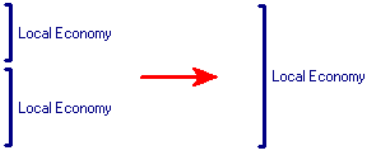
To replace all instances of the code at once, click on **Replace All**. Or, to replace one instance at a time, click on **Find Next**, and then click **Replace**.

Consolidating Coded Segments

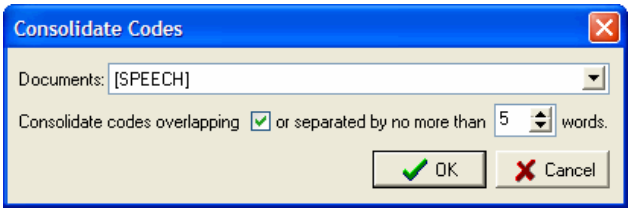
When performing extensive manual coding, autocoding and merging of codes, one may end up with some identical codes assigned more than once to the exact same text segment or to partially overlapping text segments. The CONSOLIDATE command allows one to replace codes assigned by the same coder to overlapping text segments by a single coded segment (see below).



This feature may also be used to join successive text segments sharing the same code and separated by less than a specified distance into a single, larger coded segment, like in the example below.



To consolidate overlapping or adjacent codes, run the CONSOLIDATE command from the CODES menu. The following dialog box will appear:



The **DOCUMENTS** option allows you to specify on which document variables the consolidation will be performed. If the current project contains more than one document variable, you will have a choice of selecting either one or a combination of them. By default, the currently displayed document variable is selected. To add other document variables, click on the down arrow key at the right of the list box. You will be presented with a list of all available document variables. Select the variables on which you want the consolidation to be performed.

You may choose to consolidate overlapping coded segments only or also merge coded segments that are near each other without overlapping. To consolidate such adjacent coded segments, select this option by putting a check mark in the check box and specify the maximum number of words that can appear between

the two segments. Setting this value to zero will only merge codes that appear next to each other, without any text between them. Setting this option to 5 will merge codes immediately next to each other or those that are separated by no more than five words.

To proceed with the code consolidation, click on the **OK** button. To leave the dialog box without modification to existing coded segments, click on the **Cancel** button.

Please note that code consolidation cannot be undone. If you are not entirely sure as to the overal effect of code consolidation, you should make a backup copy of your project before attempting this operation.

Coding Analysis Features

QDA Miner provides several tools to assist in the coding task and to perform descriptive, comparative and exploratory analysis of codings. These tools may be used to systematize the coding of documents, ensure coding consistency, identify regularities and patterns in coding, uncover hidden relationships between codes and other properties of the cases, etc.

The **Text Retrieval** tool searches for specific text patterns in documents. Complex search patterns may be performed using Boolean operators. This search can be performed on all documents in a project or restricted to specific ones or to some coded segments.

The **Section Retrieval** tool searches structured documents for sections bounded by fixed delimiters. This feature is especially useful to automatically assign codes to recurring sections within a document or in multiple documents. You can search in all documents in a project or restrict the search to specific document variables.

The **Keyword Retrieval** Feature can retrieve any document, paragraph, sentence, or coded segment containing a specific WordStat keyword or a combination of keywords.

The **Coding Frequency** tool allows one to obtain a list of all codes in the current codebook along with their description and various statistics such as their frequency, the number of cases in which they are found, and the total number of words in the associated text segments. This dialog box also allows one to produce bar charts and pie charts from those statistics.

The **Coding Retrieval** tool lists all text segments associated with some codes or with specific patterns of codes. Complex search patterns may include criteria such as proximity, overlapping, inclusion, sequential relationships, etc.

The **Codes Co-occurrences** tool uses information about the proximity or the co-occurrence of codes within documents to explore potential relationships among them as well as similarities among cases. This dialog box gives access to various statistical and graphic tools, such as cluster analysis, multidimensional scaling, and proximity plots.

The **Coding Sequences** tool can be used to identify recurring sequences of codes. This feature can produce frequency lists of all sequences involving two selected sets of codes as well as the percentage of time one code follows or is followed by another one.

The **Coding by Variable** tool is useful for identifying or testing potential code similarities or differences between subgroup of cases (categorical variable) or to assess the relationship between these codes and other numerical variables. Various statistical and graphical tools are used for this purpose such as contingency tables, association statistics, bar charts and line charts, heatmaps, correspondence plot, etc.

The **Inter-Coders Agreement** tool is used to compare the consistency of coding between several coders. This tool can be useful to uncover differences in interpretation, clarify equivocal rules, identify ambiguity in the text, and ultimately quantify the final level of agreement obtained by those raters.

QDA Miner may also be used in conjunction with the following two quantitative analysis software tools:

WordStat is a powerful quantitative content analysis and text-mining software. When used as an add-on to QDA Miner, it analyzes words and phrases found in specific documents or in selected code segments.

WordStat can perform simple descriptive analysis or explore in greater details the relationship between words or categories of words and other numeric or categorical variables.

Simstat is a comprehensive statistical analysis application. Since it uses the same file format as QDA Miner, Simstat can be used to perform statistical analysis on any numerical data stored in a QDA Miner project file. It can perform numeric and alphanumeric computation, transformation and recoding of variables, as well as advanced file management procedures, such as data file merging, file aggregation, etc.

Saving and Retrieving Queries

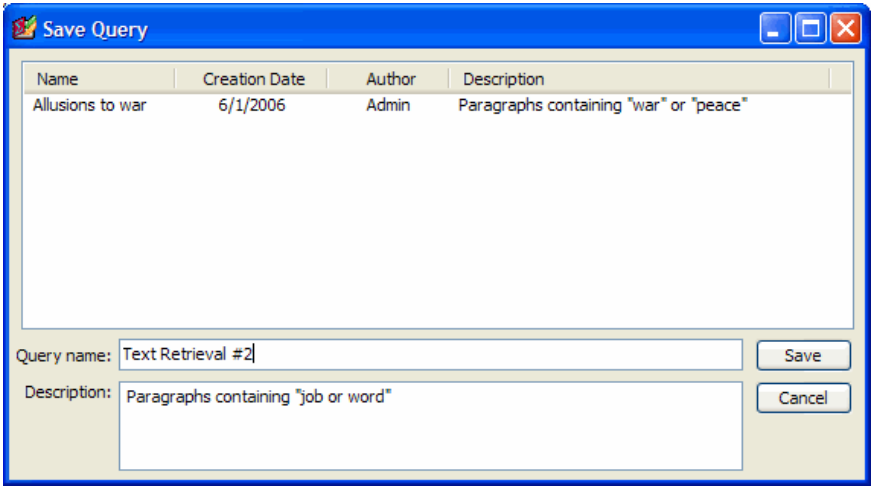
QDA Miner provides different types of queries that enable you to find specific text patterns, to select subsets of cases, and to explore the relationships among codes and between those codes and case properties. The program allows you to save with the project different kinds of queries and analysis options and retrieve them later. QDA Miner provides savings and retrievals for the following functions:

- Text Retrieval
- Section Retrieval
- Keyword Retrieval
- Coding Retrieval
- Coding Co-occurrences
- Coding Sequences
- Inter-coders Agreement
- Case Filtering

Dialog boxes associated with each of theose query types have two buttons: The  button is used to save the current question options with the project, while the  button is used to retrieve previously saved queries.

To Save a Query

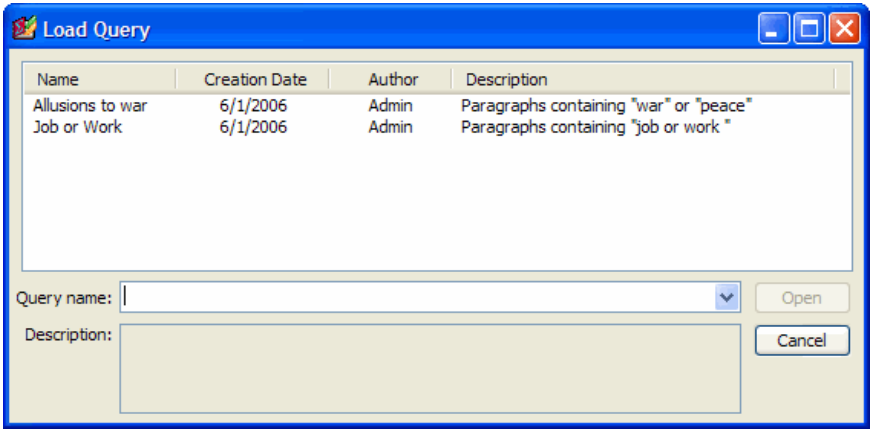
- Set the various options on the dialog box.
- Click on the  button. A dialog box similar to the one below will appear:



- The program automatically provides a generic query name and a description. Edit this name and description and click on the SAVE button to store this query with the project. If a query with an identical name already exists, you will be asked to confirm that you wish to overwrite it.

To Retrieve a Previously Saved Query

- Open the dialog box associated with the task you want to perform.
- Click on the  button. A dialog box similar to the one below will appear:



Please note that if no query of this type has been previously saved, this option will be disabled.

- From the list of previously saved queries, you can select the query you want to retrieve, either by double-clicking on its line, or click on it once and then clicking on the Open button. You may also type its name in the Query Name edit box or select it from a drop-down list and then click on the **Open** button to open it.

To delete queries

- From the **Save Query** or **Load Query** dialog box, select the query you would like to delete.
- Click on the **Del** keyboard key or right-click to display a pop-up menu and choose the **Delete** command.

To rename a query

- From the **Save Query** or **Load Query** dialog box, select the query you would like to rename.
- Right-click to display a pop-up menu and choose the **Rename** command.
- Enter the new name and click on **OK**.

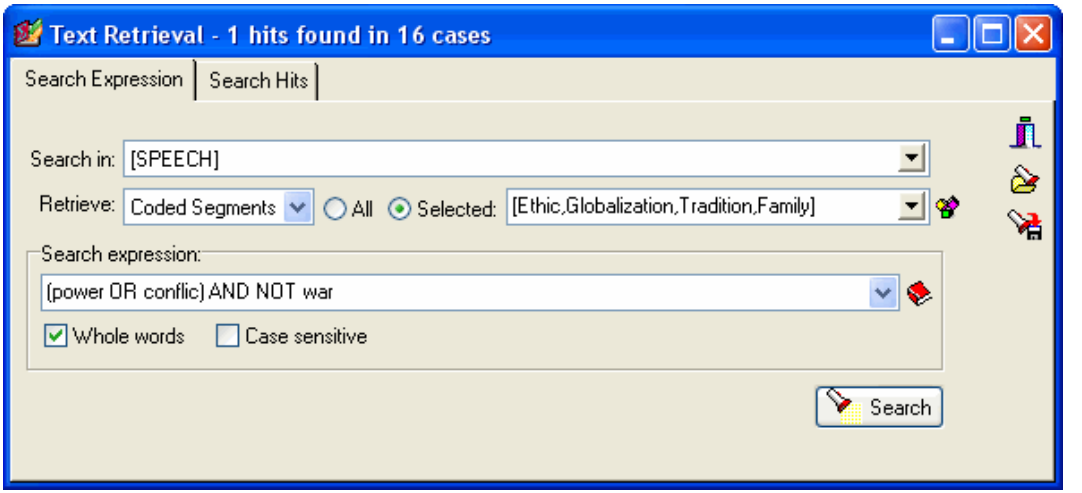
To change the description of a query

- From the **Save Query** or **Load Query** dialog box, select the query with the description you would like to modify.
- Right-click to display a pop-up menu and choose the **Edit Description** command.
- Change the description text and click on **OK**.

Text Retrieval

The TEXT RETRIEVAL function searches for specific text or combination of text in documents. You can search in all documents in a project or restrict the search to specific document variables. Searches can also be restricted to specific coded segments.

To start the text search feature, select the TEXT RETRIEVAL command from the ANALYSIS menu. The following dialog box will appear:



On the first page you can set the search conditions while results are displayed on the **Search Hits** page. When entering this dialog box, the second page is normally disabled. As soon as a search returns at least one hit, QDA Miner will enable this second page, allowing you to review the search hits and optionally assign codes to them.

The **SEARCH IN** option allows you to specify on which document variables the search will be performed. If the current project contains more than one document variable, you will have a choice of selecting either one or a combination of them. By default, all document variables are selected. To restrict the analysis to only a few of them, click on the down arrow key at the right of the list box. You will be presented with a list of all available document variables. Select the variables you wish to search.

The **RETRIEVE** option determines the search unit on which the search will be performed as well as what will be retrieved. You can select three different search units:

If you select **Documents** as the search unit, QDA Miner will apply the search expression on each document associated with a specific case. If a specific document meets the search condition, its location will be returned.

When selecting **Paragraphs** as the search unit, QDA Miner will return any paragraph meeting the search condition.

Select **Sentences** to instruct QDA Miner to return sentences meeting the search condition.

Selecting **Coded Segments** as the search unit restricts the search to the text segments that have already been coded. When a coded segment meets the search condition, its entire text is returned, no matter whether it is a single word or several paragraphs. To restrict the search to specific codes, select the **Selected** radio button on the right side of this option and select the codes from the drop-down checklist by clicking on the arrow button at the right end of this list. These codes can also be selected from a tree representation of the codebook by clicking on the  button. To perform the search on all codes, select the **All** radio button.

SEARCH EXPRESSION - This edit box ise used to specify words or phrases to search for. To search for a single word, simply enter the word in edit box. In order to search for a phrase, you need to replacethe spaces between words with underscore characters. For example, searching for the phrase "quality of life", requires you to enter "quality_of_life".

Boolean operators **AND**, **OR** and **NOT** may also be used to build complex search expressions. For example, the following search string:

angry OR mad

will retrieve any search unit containing either the word "angry" or the word "mad". Parentheses may also be used to specify the evaluation order in complex search expressions. For example, if you enter the following search expression:

(angry OR mad) AND (son OR daughter OR child)

and set the Search Unit option to Paragraphs, QDA Miner will return all paragraphs containing either "angry" or "mad", but only if this paragraph also contains either one of the three words found in the second set of parentheses (e.g. "son", "daughter" or "child").

Using wildcards in search expressions

A wildcard is a character or a set of characters that may be used in a search expression to represent one or more other characters. QDA Miner support the following wildcard:

?	Matches any single character. For example **F?T** will find **FAT, FIT** and any other combination of three characters beginning with 'F' and ending with 'T'.
*****	Matches any number of contiguous characters. For example, **T*ENT** will return every word starting with 'T' and ending with 'ENT' such as **TENT, TORRENT, TANGENT**, etc.
#	Matches any numeric character.
[ ]	Matches any of the specified characters at that position. For example, **[abc]** will match if a or b or c at that position.

[^] or **[!]** Matches anything but the specified characters at that position. For example, **[^abc]** will match any single character, but a, b, or c.

[a-e] Matches a through e at that position.

_ The underscore character may be used to replace space within words. For example searching for **QUALITITY_OF_LIFE** will match any instance of this phrase.

Performing thesaurus-based searches

A thesaurus-based search allows one to search for several words or phrases associated with a single thesaurus entry previously defined. For example, by entering a single category @GOOD (with an ampersand character "@" as a prefix), the program can automatically search for items that have been associated with this category, like "good", "fine", "excellent", "all right", "topnotch", etc. Names of categories may be typed directly in the search expression by preceeding the category with a # character. For example, the following text search expression:

@GOOD and @SERVICE

will retrieve all text units containing either one of the words or phrases associated with the thesaurus entry GOOD along with any word or phrase included in the SERVICE thesaurus entry.

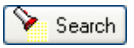
One may also insert a category, by cliking the  button to display the thesaurus editing dialog box, selecting the thesaurus entry and then clicking on the Insert button. The Thesaurus Editing dialog box allows one to create new entries or edit existing ones. For more information on thesaurus searches, see **Thesaurus Editing** on page 70.

WHOLE WORDS - This option defines whether the words in the search expressions may be part of a word or if only whole words are to be found. For example, if this option is disabled, searching for a word like "bank" will return any search unit containing not only "bank", but also "banking", "bankruptcy", etc.

CASE SENSITIVE - This option defines whether the search is case-sensitive or not. If enabled, only words with the exact same cases will be returned.

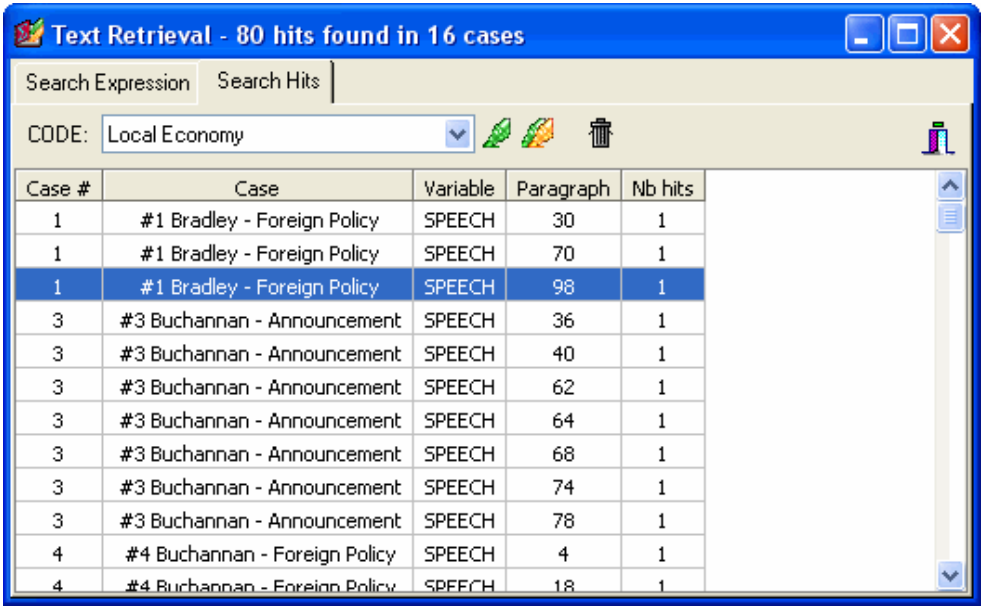
To store the search expression and options, click on the  button and specify the name under which those query options will be saved.

To retrieve a previously saved query, click on the  button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

To perform the search, click on the 

button. The results of a search are displayed in a table located on the **Search Hit** page.

The table provides basic information on each hit, such as the case number and label, as well as the document variable in which it was found, the hit location, etc. To sort this list of hits in ascending order on any column values, simply click this column header. Clicking a second time on the same column header sorts the rows again in descending order.



Case #	Case	Variable	Paragraph	Nb hits
1	#1 Bradley - Foreign Policy	SPEECH	30	1
1	#1 Bradley - Foreign Policy	SPEECH	70	1
1	#1 Bradley - Foreign Policy	SPEECH	98	1
3	#3 Buchannan - Announcement	SPEECH	36	1
3	#3 Buchannan - Announcement	SPEECH	40	1
3	#3 Buchannan - Announcement	SPEECH	62	1
3	#3 Buchannan - Announcement	SPEECH	64	1
3	#3 Buchannan - Announcement	SPEECH	68	1
3	#3 Buchannan - Announcement	SPEECH	74	1
3	#3 Buchannan - Announcement	SPEECH	78	1
4	#4 Buchannan - Foreign Policy	SPEECH	4	1
4	#4 Buchannan - Foreign Policy	SPEECH	18	1

Selecting an item in this table, either by clicking it or by using the keyboard cursor keys, such as the **UP** or **DOWN** arrow keys, automatically displays the corresponding case and document in the main window. If the search unit was set to **Paragraphs** or **Coded Segments**, selecting a specific hit will also highlight the corresponding text segment. This feature provides a way to examine the retrieved segment in its surrounding context. You can also assign an existing code to the highlighted segment.

To assign a code to a specific search hit:

- In the table of search hits, select the row corresponding to the text segment you want to code.
- Use the **CODE** drop-down list located above this table to select the code you want to assign.
- Click on the  button to assign the selected code to the highlighted text segment.

To assign a code to all search hits:

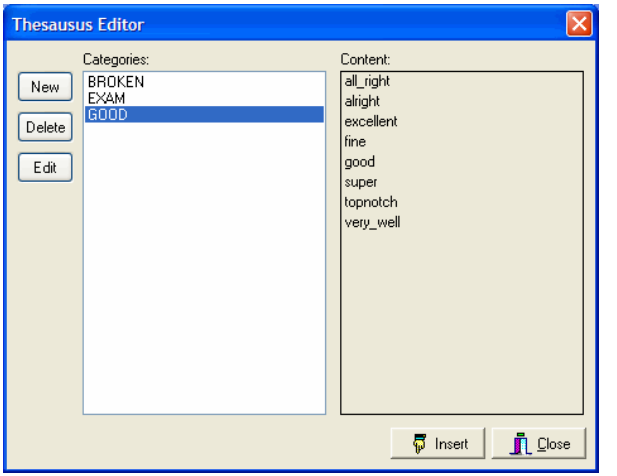
- Use the **CODE** drop-down list located above this table to select the code that you want to assign
- Click on the  button to assign the selected code to all text segments meeting the search expression.

Prior to the assignment of a code to all search hits, you may want to remove selected hits that do not correspond to what you were looking for. To remove a search hit from the list, select its row and then click on the  button.

Thesaurus Editing

A thesaurus-based text search allows one to search for several words or phrases associated with a single thesaurus entry previously defined. The Thesaurus Editor allows one to create entries that can later be used in text search expressions to retrieve text units containing anyone of those words or phrases.

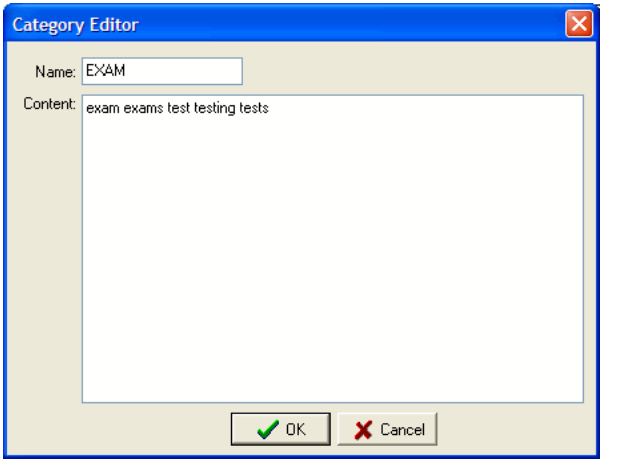
The Thesaurus Editor can be accessed from the Text Search dialog box by clicking on the icon located next to the text search expression. When activate, a dialog box similar to this one appears.



The dialog box displays a list of items currently defined. To view words and phrases associated with an entry, simply select the item in the **CATEGORIES** list. Items associated with the highlighted category are listed on the right side of the dialog box in the **CONTENT** list box.

To create an new category:

- Click on the **NEW** button. A dialog box similar to this one will appear.



- Enter the name of the category. This will be the placeholder to be used in search expression.
- Type the list of words you would like to search for when the category name is entered in a search expression. You may separate words by spaces or line breaks. Spaces in phrases should be replaced by an underscore character. For example, to search for the phrase "quality of life", you have to type "quality_of_life". You may also use any one of the valid wildcards supported by QDA Miner.
- Click on the **OK** button to store the newly defined category.

To edit an existing category:

- Select the category you would like to edit.
- Click on the **EDIT** button.
- Edit the list of items associated with the category.
- Click on the **OK** button to store the newly defined category.

To delete an existing category:

- Select the category you would like to delete.
- Click on the **DELETE** button.

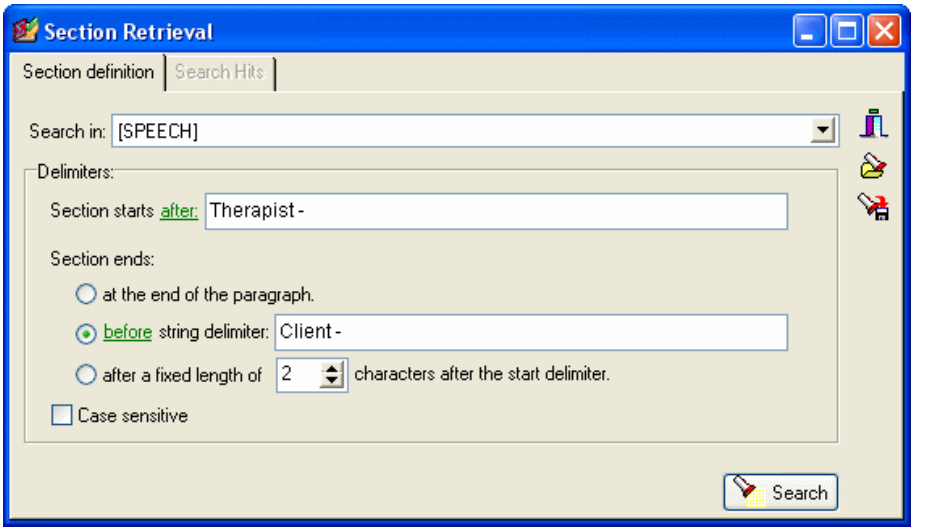
To insert a category into a search expression:

- Select the category you would like to insert.
- Click on the **INSERT** button.

Section Retrieval

The SECTION RETRIEVAL function searches structured documents for sections bounded by fixed delimiters. This feature is especially useful to automatically assign codes to recurring sections within a document or in multiple documents. You can search in all documents in a project or restrict the search to specific document variables.

To start the section search feature, select the SECTION RETRIEVAL command from the ANALYSIS menu. The following dialog box will appear:



On the first page, you can set the search conditions and results are displayed on the **Search Hits** page. When entering this dialog box, the second page is initially disabled. As soon as a search returns at least one hit, QDA Miner will enable this second page, allowing you to review the search hits and optionally assign codes to them.

The **SEARCH IN** option allows you to specify on which document variables the search will be performed. If the current project contains more than one document variable, you will have a choice of selecting either one or a combination of them. By default, all document variables are selected. To restrict the analysis to only a few of them, click on the down arrow key at the right of the list box. You will be presented with a list of all available document variables. Select the variables on which you want the search to be performed.

The **DELIMITERS** option set allows you to specify which delimiters are used to define the starting and ending location of a section.

- The **Section starts** option specifies which constant string is used to delimit the start of the section. By default, the retrieved section begins after this constant string. To include this delimiter in the retrieved segment, click on the **after** label to display a context menu and select **with**.

- The **Section ends** option allows you to select the method used to specify the length of the section. Three methods are available.

  1) If you select **at the end of the paragraph** the program retrieves all text found from the start delimiter up to end of the paragraph.

  2) Selecting **string delimiter** allows you to specific a constant string found at the end of a section. By default, the retrieved section does not include this delimiter and ends before it. To include this delimiter, simply click on the **before** label to display a context menu and select **with**.

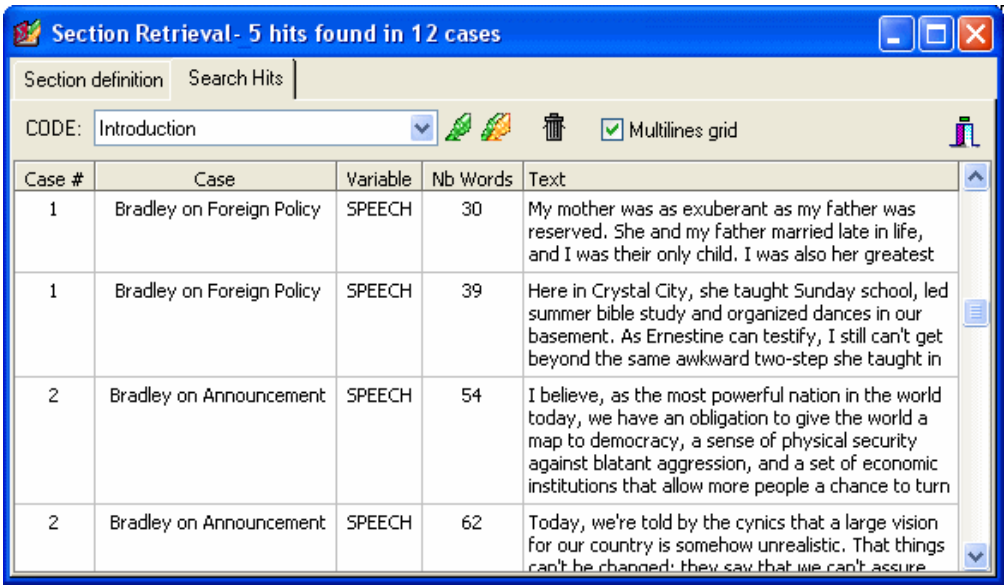
  3) The third method allows you to retrieve a fixed number of characters after the end of the starting delimiter.

- **CASE SENSITIVE** - This option defines whether the search is case-sensitive or not. If enabled, only string delimiters with the exact same cases will be recognized.

To store the search expression and options, click on the button and specify the name under which those query options will be saved.

To retrieve a previously saved query, click on the button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

To perform the search, click on the Search button. The results of a search are displayed in a table located on the **Search Hit** page.



The table provides basic information on each hit, such as the case number and label, as well as the document variable in which it was found, the hit location, etc. To sort this list of hits in ascending order on

any column values, simply click this column header. Clicking a second time on the same column header sorts the rows again in descending order.

Selecting an item in this table, either by clicking it or by using the keyboard cursor keys such as the UP or DOWN arrow keys, automatically displays the corresponding case and document in the main window. This feature provides a way to examine the retrieved segment in its surrounding context. You can also assign an existing code to the highlighted segment.

To assign a code to a retrieved section:

- In the table of search hits, select the row corresponding to the section you want to code.
- Use the CODE drop-down list locate above this table to select the code you want to assign.
- Click on the  button to assign the selected code to the highlighted section.

To assign a code to all retrieved sections:

- Use the CODE drop-down list locate above this table to select the code that you want to assign.
- Click on the  button to assign the selected code to all sections meeting the search expression.

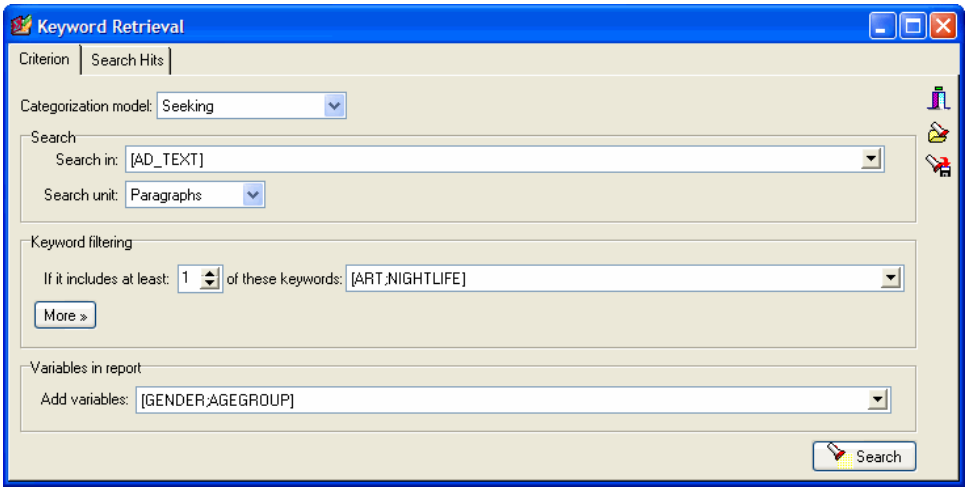
Prior to the assignment of a code to all search hits, you may want to remove selected hits that do not correspond to what you were looking for. To remove a search hit from the list, select its row and then click on the  button.

Keyword Retrieval

The KEYWORD RETRIEVAL feature can retrieve any document, paragraph, sentence, or coded segment containing a specific WordStat keyword or a combination of keywords. Text units corresponding to the search criteria are returned in a table on the Search Hits page. This table may then be printed or saved to disk. It may also be used to create tabular or text reports as well as to attach codes to the retrieved text segments.

Categorization models are advanced content analysis processes created by WordStat 5.0 or later and saved to disk in a .wcat file. A WordStat categorization model may involve various text processing such as stemming, lemmatization and word exclusions. It also involves the categorization of words, word patterns, and phrases and may also include complex coding rules involving boolean and proximity operators. Such rules may be used to perform disambiguation of words or coding of complex actions. In order to be used within QDA Miner, categorization models should be stored in a Models subfolder located under the main program folder.

QDA Miner comes with a sample categorization model developed from a content analysis of the Seeking project file, allowing users to test this feature without the need to install WordStat. This categorization model consists of 13 content categories with items such as APPEARANCE, ART, COMMUNICATION, FAMILY, etc.



The **CATEGORIZATION MODEL** option lists all WordStat categorization models available in the Models subfolder. The first step one needs to do in order to access the keyword retrieval features is selecting a categorization model. Once a model has been selected, the other search and formatting options become available.

The **SEARCH IN** option allows you to specify on which document variables the search will be performed. If the current project contains more than one document variable, you will have a choice of selecting either one or a combination of them. By default, all document variables are selected. To restrict the analysis to

only a few of them, click on the down arrow key at the right of the list box. You will be presented with a list of all available document variables. Select the variables on which you want the search to be performed.

The **SEARCH UNIT** option determines the search unit on which the search will be performed as well as what will be retrieved. You can select three different search units:

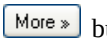
- If you select **Documents** as the search unit, QDA Miner will apply the search expression on each document associated with a specific case. If a specific document meets the search condition, its location will be returned.

- When selecting **Paragraphs** as the search unit, QDA Miner will return any paragraph meeting the search condition.

- Select **Sentences** to instruct QDA Miner to return sentences meeting the search condition.

- Selecting **Coded Segments** as the search unit allows restricts the search to the text segments that have already been coded. When a coded segment meets the search condition, its entire text is returned, no matter whether it is a single word or several paragraphs. To restrict the search to specific codes, select the **Selected** radio button to the right side of this option and select the codes from the drop-down checklist by clicking on the arrow button to the right end of this list. These codes can also be selected from a tree representation of the codebook by clicking on the  button. To perform the search on all codes, choose the **All** radio button.

KEYWORD FILTERING - This group of options allows one to select the keywords on which the retrieval will be based. Setting the numerical value of **the Include at least** option allows one to retrieve text units containing a minimum number of keywords from a selected list. To select the keywords, click on the arrow button to show all available keywords and then click on the desired items. The minimum number of keywords a specific unit must contain in order to be retrieved is specified using a small edit box with spin buttons on the right. Setting this numerical value to the total number of keywords selected will force the program to retrieve only those units containing all those keywords. Setting this number to a lower value will retrieve all text units containing at least this number of keywords from the selected list. In the example shown above, any unit containing keywords from at least one of the two categories ART and NIGHTLIFE will be retrieved.

To enter a second filtering condition, click on the 

button. You can choose to link the two filtering conditions using either one of the three Boolean expressions: AND, OR or NOT. Choosing AND will retrieve all text units fulfilling both criteria; selecting OR will result in a retrieval of text units meeting either the first or the second condition, or both, while choosing the NOT Boolean operator will retrieve text units meeting the first condition but not the second one.

To limit the filtering conditions to a single statement, click on the 

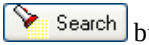
button.

ADD VARIABLES - This drop-down checklist box may optionally be used to add the values stored in one or more variables to the table of retrieved segments for the specific case from which a text segment originates. This feature is especially useful for identifying differences in how keywords are used or how

specific topics are expressed by different group of individuals. It may also be used to examine the relationship between these values and the keywords.

To store the search expression and options, click on the  button and specify the name under which those query options will be saved.

To retrieve a previously saved query, click on the  button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

Once all the search options have been set properly, simply click on the 

button to retrieve the selected text units.

Working with the Retrieved Text Units

The retrieved text units are displayed in a table found on the SEARCH HITS page. This table contains both the case number and the variable from which the segment originates, as well as the value of all additional variables selected by the user (see the ADD VARIABLES option above). When searching for paragraphs, sentences or coded segments, the table also displays the text associated with the retrieved unit and its location (its paragraph and sentence number). By using arrow keys or by clicking on a row the associated text is displayed in a separate window at the bottom of the screen with all keywords in bold.

To sort the table of retrieved units in ascending order in any column, simply click on the column header. Clicking on the same column header a second time sorts the rows in descending order. Tables may also be printed, stored as a text report, or exported to disk in various file formats such as Excel, ASCII, MS Word, HTML or XML.

To remove a search hit from the hit list:

- Select its row and then click on the  button.

To assign a code to a specific search hit:

- In the table of search hits, select the row corresponding to the text segment you want to code.
- Use the **CODE** drop-down list located above this table to select the code to assign.
- Click on the  button to assign the selected code to the highlighted text segment.

To assign a code to all search hits:

- Use the **CODE** drop-down list located above this table to select the code to assign.

- Click on the  button to assign the selected code to all text segments matching the search expression.

To create a report of retrieved segments:

- Click on the  button. The sort order of the current table is used to determine the display order in the report. This report is displayed in a text-editing dialog box and may be modified, stored on disk (in RTF, HTML or plain text format), printed, or cut-and-pasted into another application. Graphics and tables may also be inserted anywhere in this report.

To export the table to disk:

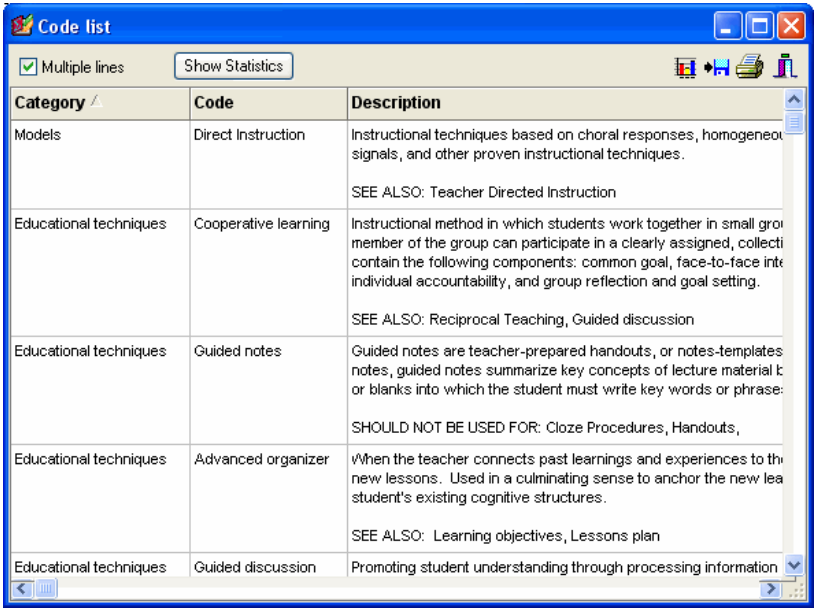
- Click on the  button. A **Save File** dialog box will appear.
- In the Save As Type list box, select the file format under which to save the table. The following formats are supported: ASCII file (*.TXT), Tab delimited file (*.TAB), Comma delimited file (*.CSV), HTML file (*.HTM; *.HTML), XML file (*.XML), MS Word document (*.Doc), and Excel spreadsheet file (*.XLS).
- Type a valid file name with the proper file extension.
- Click on the **SAVE** button.

To print the table:

- Click on the  button.

Coding Frequencies

You can use the CODING FREQUENCIES command from the ANALYSIS menu to obtain a list of all codes in the current codebook along with their description and the category to which they belong. This dialog box may also be used to obtain various statistics for each code such as their frequency, the number of cases in which they are found, and the total number of words in the associated text segments. Selecting this command displays a dialog box similar to the one below:



To spread long descriptions over more than one line, select the **Multiple lines** option.

Clicking on the **Show Statistics** button adds six columns to the right side of the table containing summary statistics on code usage:

Count	Number of times this code has been used.
Cases	Number of cases in which this code appears.
Nb Words	Total number of words in all text segments associated with this code.
% Count	Percentage of coding associated with this code.
% Cases	Percentage of cases containing this code.
% Words	Percentage of words tagged with this code.

When statistics are shown in the table, the  button allows one to produce bar charts or pie charts to visually display the distribution of specific codes. To produce such charts:

- Set the Sort By option to the order you want the values to be shown graphically.
- Select the rows you would like to plot (multiple but separate rows can be selected by clicking on the desired rows while holding down the CTRL key)
- Click on the  button.

For further information see the Bar chart and Pie Chart section below.

To remove these six columns, click on the **Hide Statistics** button.

To sort this list of hits in ascending order on any column values, simply click on a given column header. Clicking on the same column header a second time sorts the rows in descending order.

The list of codes may be saved to disk in Excel, MS Word, plain ASCII, text delimited, HTML or XML format by clicking on the  button.

You may also print this list by clicking on the  button.

Bar chart and Pie Chart

QDA Miner allows one to produce bar charts or pie charts to visually display the distribution of specific codes. To produce such charts:

- Run the CODING FREQUENCIES command from the ANALYSIS menu.
- Click on the **Show Statistics** button.
- Sort the table to the desired graphic order of the values.
- Select the rows you would like to plot (multiple but separate rows can be selected by clicking while holding down the CTRL key)
- Click on the  button.

Three types of charts may be used to depict the distribution of keywords or content categories:

The vertical bar chart is the default chart used to display absolute or relative frequencies of keywords or content categories.

The horizontal bar chart displays the same information as the vertical one but is especially useful when the number of keywords is high and their labels cannot be displayed entirely on the bottom axis.

The pie chart is useful to display the relative frequency of each keyword and compare individual values to other values and to the whole. Numerical values displayed in pie charts are always expressed in percentages of either the total frequency or case occurrences.

The **Plot** option allows one to select the values that will be used as the scale for the length of bars in bar charts or as the percentage base for pie charts. For bar charts the options are:

FREQUENCY	Number of occurrences of the code
NB CASES	Number of cases where this code appears
NB WORDS	Total number of words associated with this code
% CODES	Percentage of codes associated with this code
% CASES	Percentage of cases where this keyword appears
% WORDS	Percentage of words associated with this code

For pie charts, three options are available to specify how percentages will be computed:

FREQUENCY	Number of occurrences of the code
NB CASES	Number of cases where this code appears
NB WORDS	Total number of words associated with this code

The **View Others** option displays an additional bar or slice representing all items in the frequency table that have not been selected.

The following table provides a short description of available buttons and controls:

Controls	Description
	Press this button to retrieve a chart previously saved on disk.
	Press this button to save a chart on disk. Charts are saved in a proprietary format and may be edited and customized using the Chart Editor.
	Pressing this button allows you to print a copy of the displayed chart.
	Click on this button to turn on/off the 3D perspective for the current chart.
	This button allows you to edit various features of the chart such as the left and bottom axis , the chart and axis titles, the location of the legend, etc.

This button is used to create a copy of the chart to the clipboard. When this button is clicked, a pop-up menu appears allowing you to select whether the chart should be copied as a bitmap or as a metafile.

Pressing this button closes the chart dialog box and returns to QDA Miner's Code Frequency dialog box.

Customizing bar charts and pie charts

Clicking on the  button on the chart dialog box gives access to a dialog box to customize the appearance of bar charts and line charts. The options available in this dialog box represent only a small portion of all settings available.

To further customize the chart, modify data points, value labels, or series order, click on the  button located on the right side of the dialog box.

LEFT OR BOTTOM AXIS

Minimum / Maximum - QDA Miner automatically adjusts the vertical axis scale to fit the range of values plotted against it. To manually set these values, type the desired minimum and maximum. **Increment** Increasing or decreasing this value affects the distance between numbers as well as tick marks. Horizontal grid lines are also affected by modification of this value. **Horizontal Grid** This option turns horizontal grid lines on and off. Grid lines extend from each tick mark on an axis to the opposite side of the graph. To increase or decrease the number of grid lines or the distance between those lines, change the Increment value of the axis. A list box also allows a choice among five different line styles to draw those grid lines.

LEGEND

Location - This option positions the legend. Legends may be placed at Top, Left, Right and Bottom side of the chart.

From top - When the legend is displayed on the left or the right side of the chart, this option specifies the legend's top position in percent of total chart height.

From left - When the legend is displayed on the top or the bottom chart, this option specifies the legend's top position in percent of total chart width.

TITLES

Proper titles and axis labels are of utmost importance when describing the information displayed in a chart. By default, QDA Miner uses variable names and labels as well as other predefined settings to provide such descriptions.

The title page allows one to modify the top title, as well as the labels on the left, bottom and right axis. To edit the title, select the proper radio button. Enter several lines of text for each title by pressing the **ENTER** key at the end of a line before entering the next line.

The Font button on the right side of the edit box allows to change the font size or style of the related title.

3D VIEW

Orthogonal - Turning this option off disables the free elevation and rotation of the 3D chart.

Zoom - This option zooms the whole chart. Expressed as a percentage, increasing the value positively will bring the chart towards the viewer, increasing the overall chart size as the Zoom value increases.

3D Percent - The 3D Percent property indicates the size ratio between chart dimensions and chart depth by specifying a percent number from 1 to 100.

Perspective - Use this property with Orthogonal unchecked to modify the 3D perspective of the Chart. Larger values add more depth perspective.

Bar shadow - Enabling this option dark shades to the sides of 3D bars. Turning it off will color the sides of the bar the same as the front.

Bar width - This option determines the percent of total bar width used. Setting this value to 100 makes joined bars.

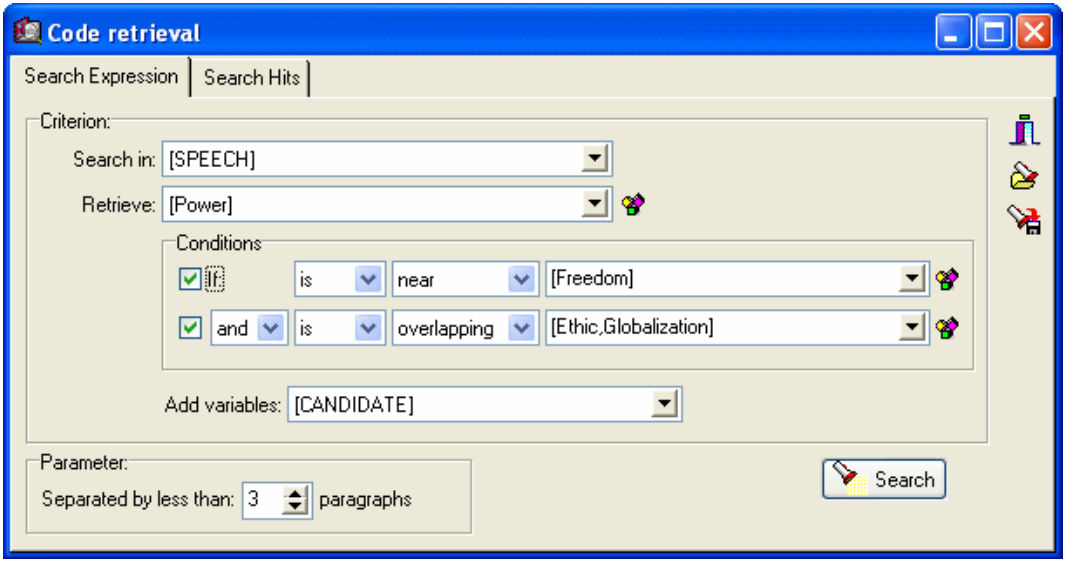
Bar depth - Use this property to limit the depth that each bar series uses. By default, bars will take up the part proportional to the number of bar series in the chart so that the back of a bar will join the front of the bar immediately behind it. To insert a gap between series of bars, decrease this value.

Pie depth - Use this property to change the thickness of the pie chart.

Coding Retrieval

The CODING RETRIEVAL feature lists all text segments associated with some codes or patterns of codes. Coded segments corresponding to the search criteria are returned in a table on the **Search Hits** page. This table may be printed or saved to disk. It may also be used to create tabular or text reports as well as to assign new codes to the retrieved segments.

To start the code retrieval feature, select the CODING RETRIEVAL command from the ANALYSIS menu. The following dialog box will appear:



Performing a Simple Code Retrieval

Only the first two options at the top of this dialog box need to be set in order to retrieve coded segments associated with some codes.

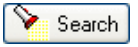
SEARCH IN - This option allows you to specify the document variables to search. If the current project contains more than one document variable, you will have a choice of selecting either one or a combination of them. By default, all document variables are selected. To restrict the search to only a few of them, click on the down arrow key at the right of the list box. You will be presented with a list of all available document variables. Select all variables on which you want the search to be performed.

RETRIEVE - This option allows you to select one or more codes to be retrieved. You can select the codes either from a drop-down checklist by clicking on the arrow button or from a tree representation of the codebook by clicking on the  button.

ADD VARIABLES - This drop-down checklist box may optionally be used to add the values stored in one or more variables to the table of retrieved segments for the specific case from which a text segment originates. This feature is especially useful for identifying differences in how codes are used or how specific topics are expressed by different group of individuals. It may also be used to examine the relationship between these values and the codes.

To store the search expression and options, click on the  button and specify the name under which those query options will be saved.

To retrieve a previously saved query, click on the  button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

Once the options have been set properly, simply click on the 

button to retrieve the selected codes.

Performing a Complex Code Retrieval

QDA Miner can also search for more complex code patterns by providing the option of setting up to two criteria specifying the spatial relationship between the codes. The combination of codes can be defined using one of seven logical operators. The table below provides the list of the available logical conditions that may be used to characterize the code combination. In this table, **Coding A** refers to the coded segments that will be retrieved while **Coding B** refers to the coded segments based on which the condition will be tested. Please note that the reverse of these functions may also be obtained by using a negation operater (see below)

Function	Description
Equal to	Coding A will be retrieved if it points to the same text segment as Coding B.
Enclosing	Coding A will be retrieved its text segment entirely enclosed another coded segment assigned to B.
Included in	Coding A will be retrieved if its text segment is entirely surrounded by another text segment assigned to B.
Overlapping	Code A will be retrieved if its text segment completely or partially overlaps another coded segment assigned to B.
Followed by	Coding A will be retrieved if followed by a text segment assigned to B. The two segments should not overlap and the distance separating the two segments should be less than a user specified distance.
Preceded by	Coding A will be retrieved if it follows a text segment assigned to B. The two segments should not overlap and the distance separating the two segments should be less than a user specified distance.
Near	Coding A will be retrieved if it follows or precedes a text segment assigned to B.

The two segments should not overlap and the distance separating the two segments should be less than a user specified distance.

Three list boxes are used to set a retrieval condition:

The **first** list box is used to specify whether to include or exclude the coded segments meeting the specified criterion. Setting this list box to **Is** will retrieve any segments meeting the specified criterion, while setting it to **Is Not** will exclude these segments and retrieve only those that do not meet the criterion.

The **second** list box allows you to choose the logical function that will be applied (see earlier for the list of logical functions). If one of the last three functions is selected (i.e., followed by, preceded by, or near) a new option will be displayed at the bottom of the dialog box. This new option allows you to set what should be the maximum distance, in number of paragraphs, separating the two codes.

The **third** list box is used to select the codes based on which the condition will be tested. If several codes are specified either in the list of codes to be retrieved or in this third list box, they will be treated as if they were joined by the **OR** Boolean operator. In other words, the condition will be met if any combination of codes is encountered. For example, searching for codes A and B to the condition that they must be equal the list of codes C, D and E corresponds to the following formula:

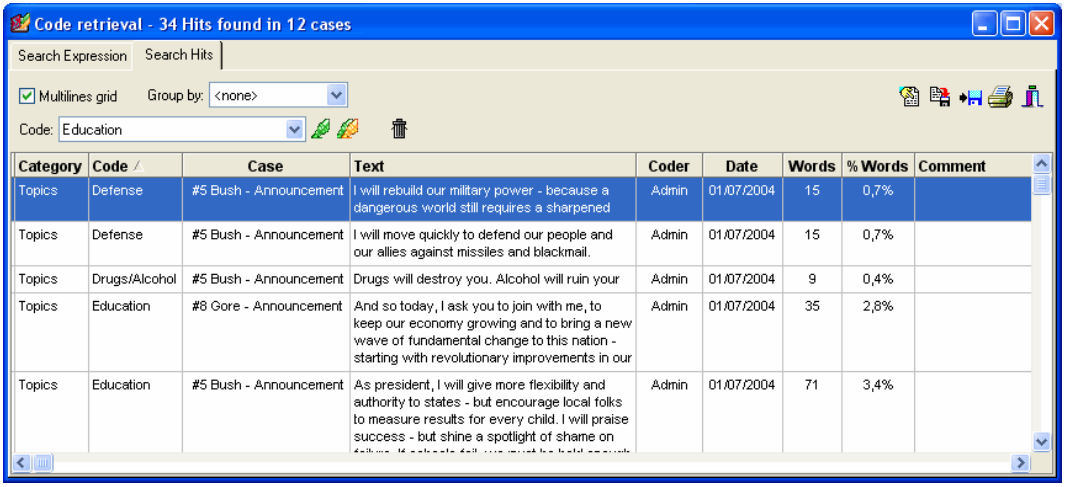
(A or B) is equal to (C or D or E)

Up to two retrieval conditions may be set and combined using a Boolean operator. If the **AND** operator is used, then only these segments meeting both criteria will be retrieved while selecting the **OR** operator will retrieve any segment meeting at least one of these two criteria.

For example, in the dialog box above, QDA Miner will retrieve any text segments coded as *Power* if this segment is within three paragraphs of another text segment coded as *Freedom*, and if it overlaps a text segment coded either as *Ethic* or *Globalization*.

Working with the Retrieved Segments

The retrieved codes are displayed in a table found on the **Search Hits** page.



Category	Code	Case	Text	Coder	Date	Words	% Words	Comment
Topics	Defense	#5 Bush - Announcement	I will rebuild our military power - because a dangerous world still requires a sharpened	Admin	01/07/2004	15	0,7%	
Topics	Defense	#5 Bush - Announcement	I will move quickly to defend our people and our allies against missiles and blackmail.	Admin	01/07/2004	15	0,7%	
Topics	Drugs/Alcohol	#5 Bush - Announcement	Drugs will destroy you. Alcohol will ruin your	Admin	01/07/2004	9	0,4%	
Topics	Education	#8 Gore - Announcement	And so today, I ask you to join with me, to keep our economy growing and to bring a new wave of fundamental change to this nation - starting with revolutionary improvements in our	Admin	01/07/2004	35	2,8%	
Topics	Education	#5 Bush - Announcement	As president, I will give more flexibility and authority to states - but encourage local folks to measure results for every child. I will praise success - but shine a spotlight of shame on	Admin	01/07/2004	71	3,4%	

Beside the text associated with retrieved coded segments, this table also contains the code name along with its category, the case from which this segment originates, the coder name and the date when it has been assigned. The **Words** column contains the total number of words in the text segment, while the **%Words** column shows the relative importance of this segment versus the entire document, in term of percentage of words. If a comment has been assigned to a specific segment, it will also be included in the table. Lastly, values for all variables that you have chosen to add to the table will be listed at the right end of the table.

To sort this list of retrieved segments in ascending order on any column values, simply click this column header. Clicking on the same column header a second time sorts the rows again in descending order.

Hits tables may be printed, coded, stored as a text report, saved to disk either as a new project or exported in a new file format such as Excel, MS Word, ASCII, HTML, or XML format.

To remove a search hit from the hit list:

- Select its row and then click on the  button.

To assign a code to a specific search hit:

- In the table of search hits, select the row corresponding to the text segment you want to code.
- Use the **Code** drop-down list located above this table to select the code that you want to assign.
- Click on the  button to assign the selected code to the highlighted text segment.

To assign a code to all search hits:

- Use the **Code** drop-down list located above this table to select the code that you want to assign.
- Click on the  button to assigned the selected code to all text segments meeting the search expression.

To create a report of coded segments:

- Click on the  button. The sort order of the current table is used to determine the display order in the report. This report is displayed in a text editor dialog box and may be modified, stored on disk in RTF, HTML or plain text format, printed, or cut and pasted into another application. Graphics and tables may also be inserted anywhere in this report.

To export retrieved codes to disk:

- Click on the  button. A **Save File** dialog box will appear.
- In the **Save as type** list box select the file format under which you would like to save the table. The following formats are supported: ASCII file (*.TXT); Tab delimited file (*.TAB); Comma delimited file (*.CSV); HTML file (*.HTM; *.HTML); XML file (*.XML), MS Word document (*.doc) Excel spreadsheet file (*.XLS).
- Enter a valid filename with the proper file extension.
- Click on the **Save** button.

The obtained table can also be stored in a new project file, where each row becomes a new case and each column is transformed into a variable. This type of project file may be useful for performing a more detailed analysis of coded segments.

To export retrieved codes into a new project:

- Click on the  button.
- Enter the name of the new project and click on the **Save** button.

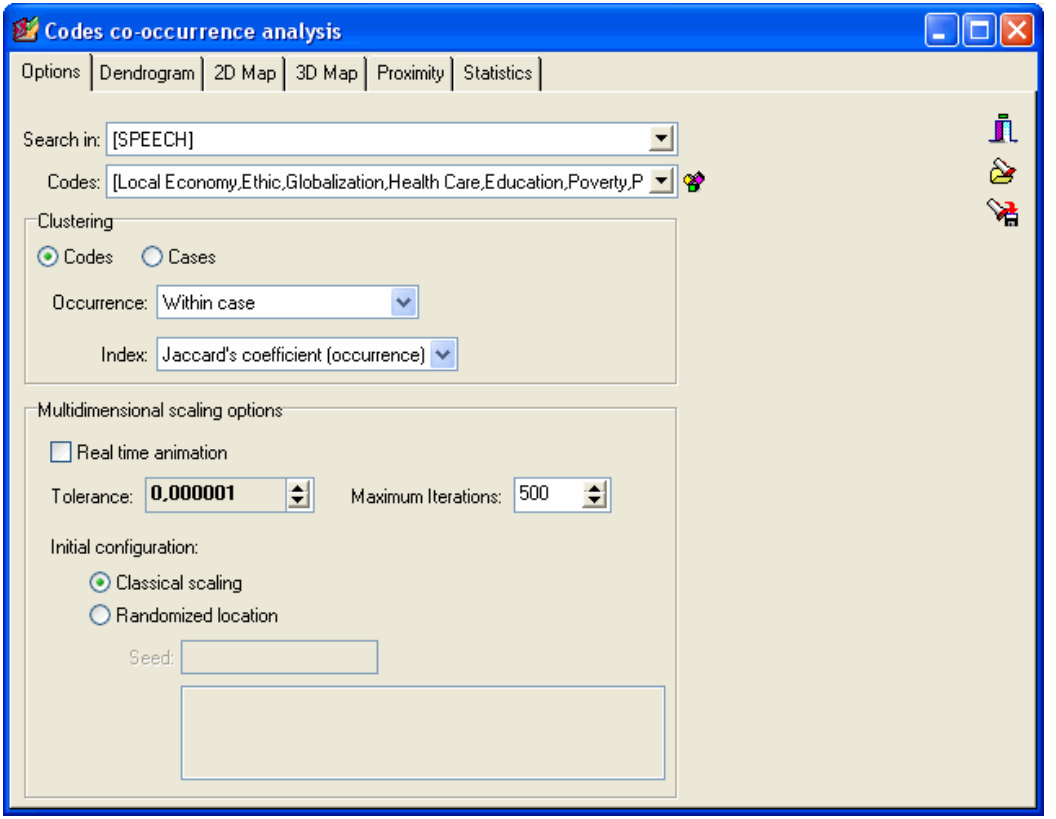
To print the table:

- Click on the  button.

Coding Co-occurrences

QDA Miner allows you to further explore the relationship among codes by providing various graphic tools to assist the identification of related codes. These tools are obtained through the computation of similarity or co-occurrences index and the application of hierarchical cluster analysis and multidimensional scaling on all or selected codes. Results are displayed in the form of dendrograms, concept maps and proximity plots. Cases may also be clustered based on their content similarity using the same statistical and graphic tools.

To perform this type of analysis, select the CODING CO-OCCURRENCES command from the ANALYSIS menu. A dialog box similar to the one below will appear:



The first page of the dialog box is used to restrict the analysis to specific document variables or specific codes. It specifies whether clustering should be performed on codes or on cases and can be used to set various analysis and display options for both enters of analysis.

SEARCH IN - This option allows you to specify on which document variables the analysis will be performed. If the current project contains more than one document variable, you will have a choice of selecting either one of them or a combination of them. By default, all document

variables are selected. To restrict the analysis to a few of them, click on the down arrow key at the right of the list box. You will be presented with a list of all available document variables. Select all variables on which you want the search to be performed.

CODES - This option allows you to select which codes will be analyzed. By default, the analysis is performed on all codes in the codebook. You can select specific codes either from a drop-down checklist by clicking on the arrow button or from a tree representation of the codebook by clicking on the  button.

Clustering Cases

When the clustering is set to be performed on cases, the distance matrix used for clustering and multidimensional scaling consists of cosine coefficients computed on the relative frequency of the various codes. The more similar two cases will be in term of the distribution of codes, the higher will be this coefficient.

DATA - This option allows you to select which indicator will be used to compare the importance of codes in the various cases. Four indicators are available:

- Code Occurrence
- Code Frequencies
- Number of Words
- Percentage of Words

Selecting **Code Occurrence** will result in a comparison based on which codes appear in each case, without taking into account the number of times each code appears. To take into account the number of times a code has been used, select **Code Frequencies**. Selecting **Number of Words** will base the comparison on the total number of words that have been assigned to different codes, while the **Percentage of Words** option will divide this value by the total number of words found in documents associated with each case.

Clustering Codes

When clustering codes, several options are available to define co-occurrence and select which similarity index will be computed from the observed co-occurrences.

CO-OCCURRENCE - This option allows you to specify how a co-occurrence will be defined. By default, a co-occurrence is said to happen every time two codes appear in the same document (within document option). You may also restrict the definition of co-occurrence to codes that are separated by no more than a specified number of paragraphs (window of n paragraphs option), or to codes that overlap each other (segment overlap option). Lastly, you can also restrict the definition of co-occurrence even further by limiting it to instances where codes have been assigned to the exact same segment.

INDEX - The Index option allows the selection of the similarity measure used in clustering and in multidimensional scaling. Four measures are available. The first three measures are based on the mere occurrences of specific codes in a case and do not take into account their frequency. In all these indexes, joint absences are excluded.

Jaccard's coefficient - This coefficient is calculated from a fourfold table as *a/(a+b+c)*, where *a* represents cases where both items occur, and *b* and *c* represent cases where one item is found but not the other. In this coefficient equal weight is given to matches and non matches.

Sorensen's coefficient - This coefficient is similar to Jaccard's but matches are weighted double. Its formula is *2a/(2a+b+c)*, where *a* represents cases where both items occur, and *b* and *c* represent cases where one item is present but the other one is absent.

Ochiai's coefficient - This index is the binary form of the cosine measure. Its formula is *SQRT(a^2/((a+b)(a+c)))*, where *a* represents cases where both items occur, and *b* and *c* represent cases where one item is present but not the other one.

The last coefficient takes into account not only the presence of a code in a case, but also how often it appears in this case.

Cosine theta - This coefficient measures the cosine of the angle between two vectors of values. It ranges from -1 to +1.

The options below apply whether clustering cases or codes. They will affect either the calculation or the display of multidimensional scaling charts.

REAL-TIME ANIMATION - When this option is enabled, the multidimensional plots are updated after every iteration allowing the user to monitor the progress made during the analysis at the cost of higher computing time.

TOLERANCE - This option specifies the tolerance factor that is used to determine when the algorithm has converged with a solution. Reducing the tolerance value may produce a slightly more accurate result, but will increase the number of iterations and the running time.

MAXIMUM ITERATIONS - This option allows you to specify the maximum number of iterations that are to be performed during the fitting procedure. If the solution does not converge with the limit specified by the TOLERANCE option before the maximum number of iterations is reached, the process is stopped and the results are displayed.

INITIAL CONFIGURATION - This option allows you to specify whether multidimensional scaling will be applied on a random configuration of points or on the result of classical scaling.

By selecting the **Classical Scaling** option, you perform a classical scaling first on the similarity matrix, and then use the derived configuration as initial values for the ordinal multidimensional scaling analysis.

Selecting the **Randomized Location** option performs multidimensional scaling analysis on a random configuration of points. By default, QDA Miner initializes the random routine before each analysis with a new random value. The seed value used for the creation of this initial configuration is stored along with the final stress value in the **history** listbox, located at the bottom of the dialog box. The **Seed** option may be used to specify a starting number that will initialize the randomization process and produce a fixed random sequence. To recall a specific seed value used previously, double-click on the proper line in the **history** list box.

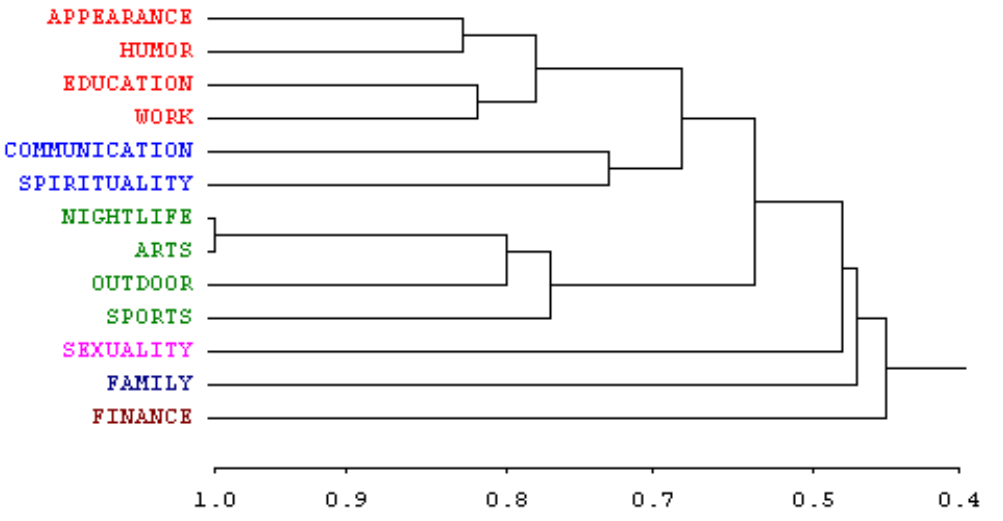
To store the analysis options, click on the  button and specify the name under which those query options will be saved.

To retrieve a previously saved options, click on the  button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

Once the various options have been set, move to the appropriate page to perform the requested analysis.

Dendrogram

QDA Miner uses an average-linkage hierarchical clustering method to create clusters from a similarity matrix. The result is presented in the form of a dendrogram (see below), also know as a tree graph. In this type of graph, the vertical axis is made up of the items, and the horizontal axis represents the clusters formed at each step of the clustering procedure. Codes that tend to appear together are combined at an early stage, while those that are independent from one another or those that do not appear together tend to be combined at the end of the agglomeration process.



It is important to remember that a dendrogram only specifies the temporal order of the branching sequence. Any cluster can be rotated around each internal branch on the tree without in any way affecting the accuracy of the dendrogram. The best analogy here is to think of a Calder mobile. Different photos for this type of mobile will yield different distances between hanging objects.

The following controls are found at the top of the dendrogram page:

Nb Clusters This option sets how many clusters the clustering solution should have. Different colors are used both in the dendrogram and in the 2D and 3D maps to indicate membership of specific items to different clusters.

Display This option lets you choose whether the vertical lines of the dendrogram represents the agglomeration schedule or the similarity indices.

 Use this button to increase the dendrogram font size and focus on a smaller portion of the tree.

 Use this button to reduce the dendrogram font size and view a larger portion of the tree.

 This button allows storing the displayed dendrogram into a graphic file. QDA Miner supports three different file formats: .BMP (Windows bitmap files), .PNG (Portable Network Graphic compress files) and .JPG (JPeg compressed files).

Note: Clustering using other similarity or distance measures or agglomeration methods may be achieved using Simstat and the MVSP multivariate analysis add-on module.

2D and 3D Concept Maps

The concept maps are graphic representations of the proximity values calculated on all included items (codes or cases) using multidimensional scaling (MDS). When clustering codes, each point represents a code and the distances between pairs of points indicate how likely these codes tend to appear together. In other words, codes that appear close together on the plot usually tend to occur together, while codes that are independent from one other or that do not appear together are located on the chart far from each other. The interpretation of multidimensional scaling plot is somewhat different when analyzing cases. In this case, each point represents a case and the distance between pairs of points indicate how similar the two cases are. Cases with similar patterns of codes will tend to appear close each other, while dissimilar cases will be plotted far from each other. Colors are used to represent membership of specific items to different partitions created using hierarchical clustering. The resulting maps are useful to detect meaningful underlying dimensions that may explain observed similarities between items.

Please note that since multidimensional scaling attempts to represent the various points into a two- or three-dimensional space, some distortion may result, especially when this analysis is performed on a large number

of items. As a consequence, some items that tend to appear together or that are parts of the same cluster may still be plotted far from each other. As well, performing a multidimensional scaling on a large number of items usually produces a cluttered map that is hard to interpret. For these reasons, interpretation of concept maps may be feasible only when applied to a relatively limited number of items.

2D and 3D Map controls

NO. CLUSTERS - This option is used to set how many clusters the clustering solution should have. Different colors are used both in the dendrogram and in the 2D and 3D maps to indicate membership of specific items to different clusters.

The actual orientation of axes in the final solution is arbitrary. The map may be rotated in any way desired; the distances between items remain the same. The rotating knob can be used to adjust the final orientation of axes in the plane or space in order to obtain an orientation that can be most easily interpreted.

Clicking this button enables you to zoom in on a plot. To zoom in on a area of the plot, hold the left mouse button down and drag the mouse down and to the right. A rectangle indicates the selected area. Release the left mouse button to zoom in.

Clicking this button restores the original viewing area of the plot.

Clicking this button performs another multidimensional scaling on a new random configuration of points. This button is visible only when the initial configuration is set to **Random Location**.

This button is used to create a copy of the chart to the clipboard. When this button is clicked, a shortcut menu appears in which you can select whether the chart should be copied as a bitmap or as a metafile.

This button allows editing of various features of multidimensional scaling plots such as the appearance of value labels and data points, chart and axes titles, the location of the legend, etc. (see Multidimensional Scaling Plot Options below)

This button allows storing the displayed multidimensional scaling plot in a graphic file. QDA Miner supports four different file formats: .BMP (Windows bitmap files), .PNG (Portable Network Graphic compress files) and .JPG (JPeg compressed files) as well as .WSX a proprietary file format (QDA Miner Chart file). Charts stored in the latter format may be opened, further edited and customized using the external Chart Editor utility program.

Clicking this button prints a copy of the displayed chart.

Clicking this button closes the **Dendrogram & Concept Map** dialog box and returns to QDA Miner's main window.

3D Map buttons

This button shows or hides left, bottom and back walls.

Clicking this button draws anchor lines from the floor to the data point to better locate data points in all 3 dimensions.

Clicking this button swithches the data of the X, Y and Z axes.

Clicking this button allows changing the viewing angle of the chart. To rotate the chart, make sure this button is selected, click any area of the chart, hold the mouse button down and drag the mouse to apply the desired rotation.

Locating a data point on the depth dimension of a 3D plot can be very difficult, especially when the plot remains static. One often has to rotate this plot constantly on the various axes to obtain an accurate idea of where the data point is located on this third axis. Clicking this button forces QDA Miner to rotate the plot automatically. To disable automatic rotation, click a second time.

Multidimensional Scaling Plot Options Dialog Box

The various options in this dialog box are used to customize the appearance of multidimensional scaling plots. These options represent only a small portion of all settings available.

To further customize the chart, edit data points, value labels, etc., click on the  button located at the right side of the dialog box.

DATA POINTS

VIEW DATA POINTS - By default, multidimensional scaling plots display both the data points and their associated labels. This option toggles on and off the display of the data points.

STYLE - This option defines the shape used to display the data points. Nine different pointer shapes may be used to represent data point (e.g. square, triangles, dots, cross, etc.).

SHADOW - This option toggles on and off the shade on the sides of data points. This option only affects square, triangle and down triangle data points when viewed in three-dimension.

SIZE - This option specifies the width and height of data points in pixels.

DATA LABELS

TRANSPARENT LABELS - This option allows you to specify whether data labels are to be displayed on an opaque background or whether the background should be invisible.

BACKGROUND - This option allows you to select the background color of data labels when not transparent

BORDER - This option enables or disables the rectangle surrounding the label background.

VERTICAL GAP - The vertical gap property determines the vertical distance in pixels between the top of the data point and the bottom of its label.

FONT - Click this button to adjust the font properties used to display data labels (i.e., style, size, and color).

WALLS AND FRAME PAGE

WALL VISIBLE - Use this option to toggle the display of left, back and bottom walls to simulate a 3D effect.

TRANSPARENT - The Transparent property controls whether the wall backgrounds will be opaque or transparent.

DARK 3D - This option shades the sides of the walls.

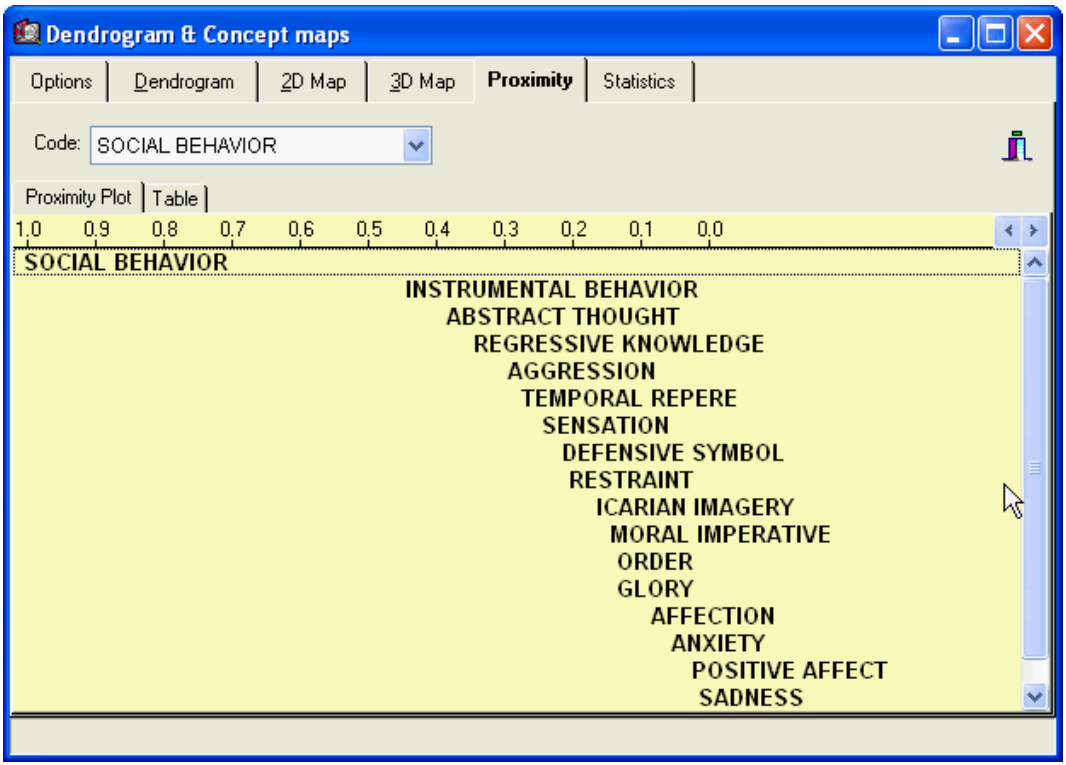
TITLE

SHOW TITLE - By default, multidimensional scaling plots have no title. To display a title, enable the Show Title option and enter the desired title in the edit box to the right of this check box. Enter several lines of text by pressing the **<Enter>** key at the end of a line before entering the next line.

The **FONT** button at the bottom of this page is used to change the font size or style of this title.

Proximity Plot

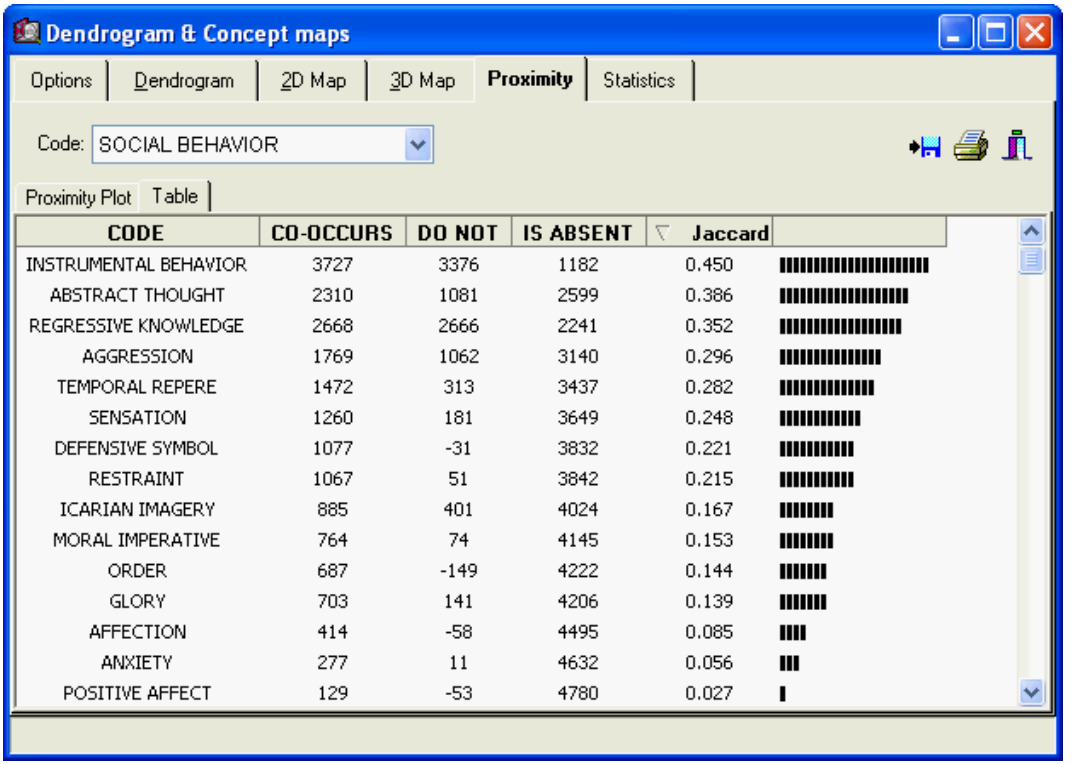
Cluster analysis and multidimensional scaling are both data reduction techniques and may not accurately represent the true proximity of codes or cases to each other. In a dendrogram, while codes that co-occur or cases that are similar tend to appear near each other, one cannot really look at the sequence of codes as a linear representation of these distances. You have to remember that a dendrogram only specifies the temporal order of the branching sequence. Consequently, any cluster can be rotated around each internal branch on the tree without in any way affecting the accuracy of the dendrogram. The best analogy here is to think of a Calder mobile. Different photos for this type of mobile will yield different distances between hanging objects. While multidimensional scaling is a more accurate representation of the distance between objects, the fact that it attempts to represent the various points in a two- or three-dimensional space may result in distortion. As a consequence, some items that tend to appear together or to be very similar may still be plotted far from each other.



The proximity plot is the most accurate way to graphically represent the distance between objects by displaying the measured distance from a selected object to all other objects on a single axis. It is not a data reduction technique but a visualization tool to help you extract information from the huge amount of data stored in the distance matrix at the origin of the dendrogram and the multidimensional scaling plots. In this plot, all measured distances are represented by the distance from the left edge of the plot. The closer an object is to the selected one, the closer it will be to the left.

To select a code or a case that will be used as the point of reference, you can choose it from the CODES or CASE drop-down list located at the top of the page. You can also freely browse through different codes or cases by double-clicking its line in the proximity plot.

The table view allows you to examine in more detail the numerical values behind the calculation of these plots. When the distance measure is based on co-occurrences, the table provides detailed information such as the number of times a given code co-occurs with the selected one (CO-OCCURS) and the number of times that it appears in the absence of this selected code (DO NOT). This table also includes the number of times the selected code appears in the absence of the given code (IS ABSENT). In the example below, we can see on the first line of the table that the INSTRUMENTAL BEHAVIOR code co-occurs 3727 times with SOCIAL BEHAVIOR, but on 3376 occasions, this code is encountered without the presence of SOCIAL BEHAVIOR, while this last code is found to occur 1182 times in the absence of INSTRUMENTAL BEHAVIOR.

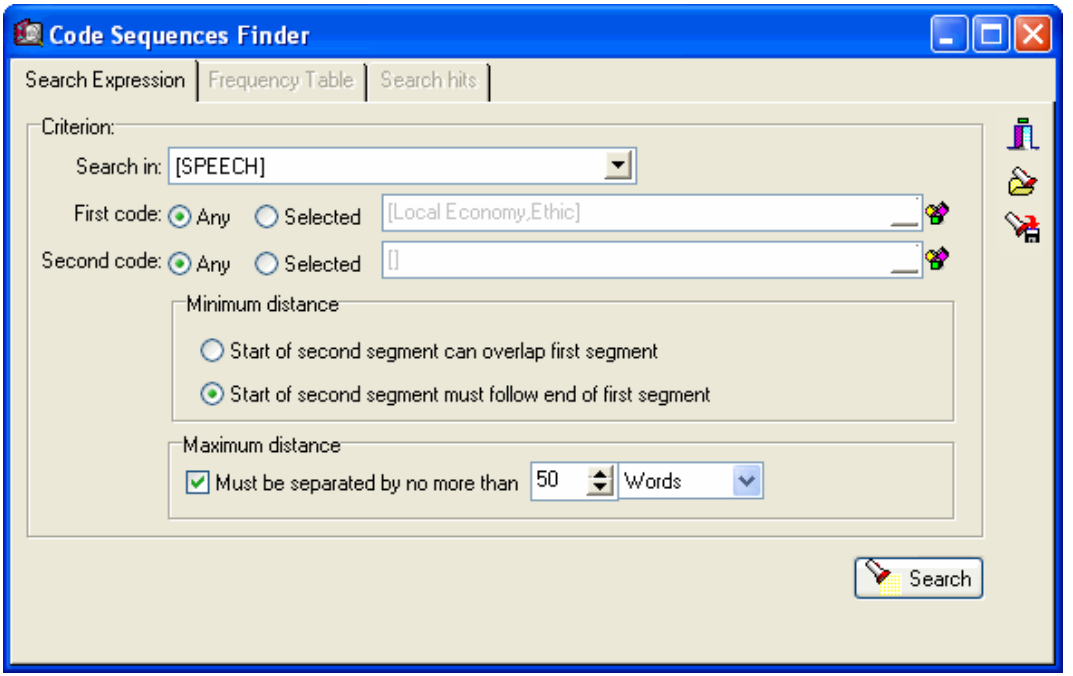


CODE	CO-OCCURS	DO NOT	IS ABSENT	Jaccard
INSTRUMENTAL BEHAVIOR	3727	3376	1182	0.450
ABSTRACT THOUGHT	2310	1081	2599	0.386
REGRESSIVE KNOWLEDGE	2668	2666	2241	0.352
AGGRESSION	1769	1062	3140	0.296
TEMPORAL REPERE	1472	313	3437	0.282
SENSATION	1260	181	3649	0.248
DEFENSIVE SYMBOL	1077	-31	3832	0.221
RESTRAINT	1067	51	3842	0.215
ICARIAN IMAGERY	885	401	4024	0.167
MORAL IMPERATIVE	764	74	4145	0.153
ORDER	687	-149	4222	0.144
GLORY	703	141	4206	0.139
AFFECTION	414	-58	4495	0.085
ANXIETY	277	11	4632	0.056
POSITIVE AFFECT	129	-53	4780	0.027

Code Sequence Analysis

While coding co-occurrences analysis looks at the concomitant presence of codes regardless of their order of appearance in the documents, the Code Sequences finder is used to identify codes that not only co-occur, but do so in a specific order and under specific conditions. This command list the frequency of all sequences involving two selected sets of codes as well as the percentage of time one code follows or is followed by another one.

To search for code sequences, select the CODING SEQUENCES analysis from the ANALYSIS menu. A dialog box similar to the one below will appear:



The first page of the dialog box is used to restrict the analysis to specific document variables or specific codes and to set the conditions that must be met to determine whether a code follows another one.

SEARCH IN - This option allows you to specify the document variable to search. If the current project contains more than one document variable, you will have a choice of selecting either one or more of them. By default, all document variables are selected. To restrict the analysis to a few of them, click on the down arrow key at the right of the list box. A list of all available document variables will appear. Select all variables on which you want the search to be performed.

FIRST CODES and **LAST CODES** - These two options allow you to select which specific code sequences will be analyzed. By default, the analysis is performed on all codes in the codebook, providing you with a comprehensive list of all sequences found in the documents. The list of sequences can however be limited by restricting either the leading or the following codes. For example, if someone is interested in the reactions of a therapist to client verbalizations, he may set the **First Codes** option to include only the codes used to describe client verbalizations and limit the codes in the **Last Codes** option to those used to describe the therapist's reaction. You can select specific codes for these two options either from the corresponding drop-down checklist by clicking on the arrow button or from a tree representation of the codebook by clicking on the  button.

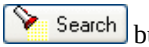
MINIMUM DISTANCE - This option is used to specify whether two codes will be considered to form a sequence when they partially overlap. If the first option is selected, a code can be considered as following the first code despite the fact that the text segment to which it was assigned starts before the end of the segment associated with the first code. However, if the second radio button is selected, QDA Miner will not consider these overlapping codes to be code sequences. For example, in the following example:

If the option is set to **Start of second segment can overlap the first segment**, then QDA Miner will treat the *Local Economy* and *Free Trade* codes as a valid sequence. If the **Minimum Distance** option is set so that the start of the second code must follow the end of the first one, then the same pair of codes will not be considered as a sequence and will not be included in the results table.

MAXIMUM DISTANCE - This option is used to specify what should be the maximum distance separating two codes in order to consider them to form a sequence. If disabled, QDA Miner will consider a code to follow another one as long as it is in the same document and meets the minimum distance criteria (see above). When enabled, code sequences are restricted to codes separated by no more than a predefined distance. This distance is calculated from the end of the first coded segment to the beginning of the second coded segment and can be measured either in number of characters, words, or paragraphs, or alternatively by setting a maximum number of codes that can occur between them. If the minimum distance is set to zero code, then only the codes following immediately the first code will be considered to be a sequence.

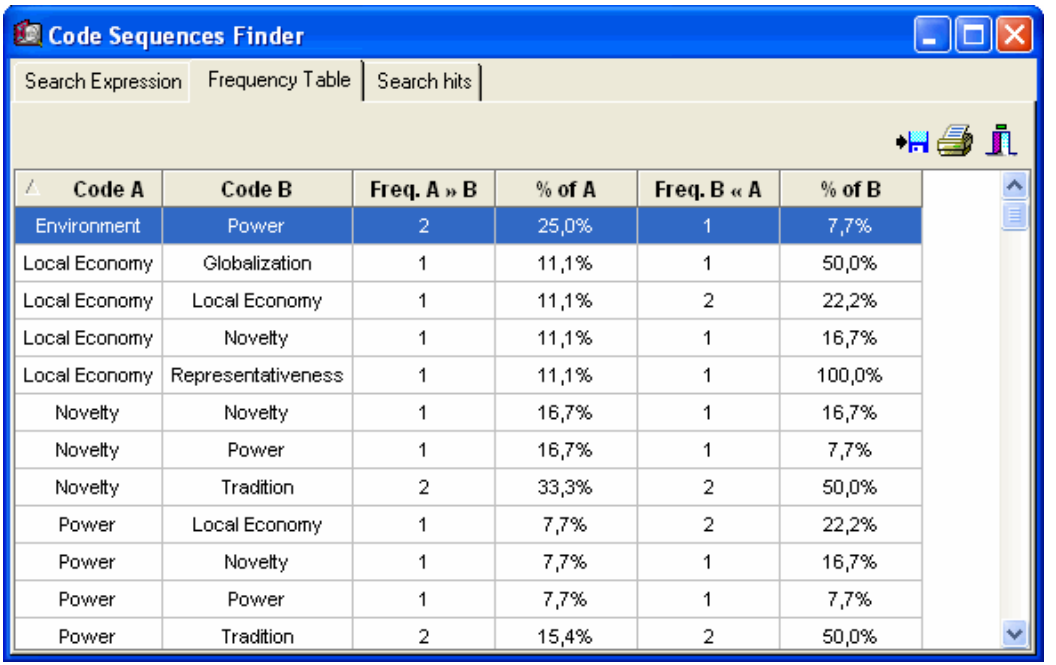
To store the search expression and options, click on the  button and specify the name under which those query options will be saved.

To retrieve a previously saved query, click on the  button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

Once the criteria have been set properly, click on the 

button to find all sequences meeting those criteria. If at least one sequence if found, the **Frequency Table** and **Search Hits** page become enabled.

The **Frequency Table** page shows a frequency list of all code sequences found in the document.



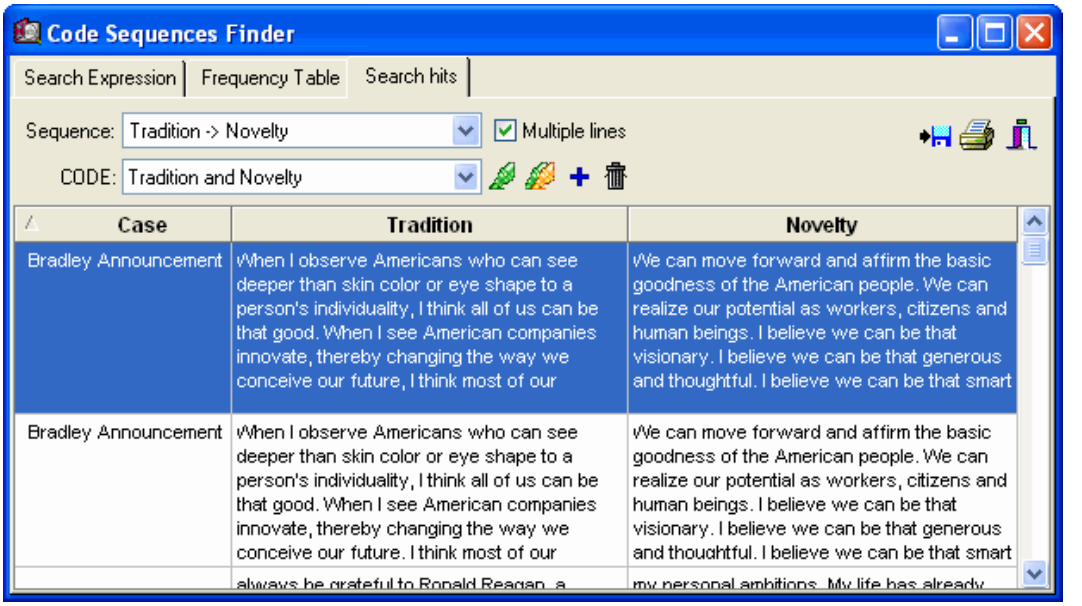
Code A	Code B	Freq. A » B	% of A	Freq. B « A	% of B
Environment	Power	2	25,0%	1	7,7%
Local Economy	Globalization	1	11,1%	1	50,0%
Local Economy	Local Economy	1	11,1%	2	22,2%
Local Economy	Novelty	1	11,1%	1	16,7%
Local Economy	Representativeness	1	11,1%	1	100,0%
Novelty	Novelty	1	16,7%	1	16,7%
Novelty	Power	1	16,7%	1	7,7%
Novelty	Tradition	2	33,3%	2	50,0%
Power	Local Economy	1	7,7%	2	22,2%
Power	Novelty	1	7,7%	1	16,7%
Power	Power	1	7,7%	1	7,7%
Power	Tradition	2	15,4%	2	50,0%

The table includes the following information:

Column	Interpretation
Code A	The name of the first code in the sequence
Code B	The name of the second code in the sequence
Freq. A » B	The number of times that the first code is followed by this second code.
% of A	The percentage of time that the first code is followed by this second code.
Freq. B « A	The number of times that the second code follows this first code.
% of B	The percentage of time that the second code follows this first code.

Clicking any column header sorts the rows in ascending order of values found in this column. Clicking on the same column header a second time sorts the table in descending order.

The **Search Hits** page allows you to examine in more detail the text segments associated with specific code sequences.



To select the code sequence to display, select its name from the **Sequence** list box. Selecting an item in this table, either by clicking its row or by using keyboard cursor keys such as the **UP** or **DOWN** arrow keys automatically displays the corresponding case and document in the main window. This feature provides a way to examine the retrieved segments in their surrounding context.

To remove a search hit from the hit list:

- Select its row and then click on the  button.

To assign a code to a specific search hit:

- In the table of search hits, select the row corresponding to the text segment you want to code.
- Use the **CODE** drop-down list located above this table to select the code to assign.
- Click on the  button to assign the selected code to the highlighted text segment.

To assign a code to all search hits:

- Use the **CODE** drop-down list located above this table to select the code to assign.

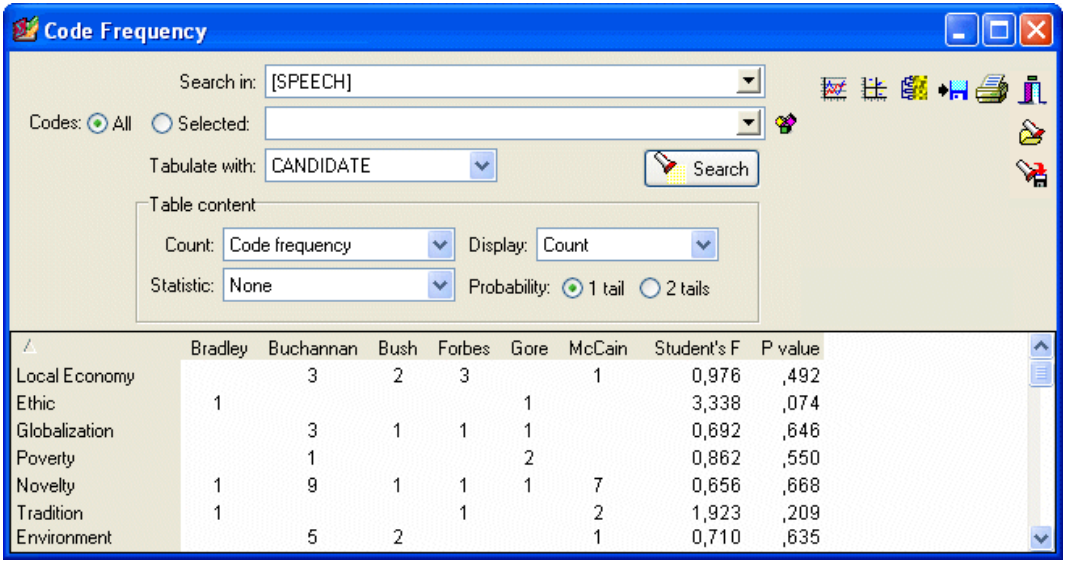
- Click on the  button to assign the selected code to all text segments matching the search expression.

The **Frequency Table** or the **Search Hits** list may be saved to disk in Excel, MS Word, plain ASCII, text delimited, HTML, or XML format by clicking on the  button.

To print the table, click on the  button.

Comparing Coding by Variables

The CODE FREQUENCY dialog box is used to explore the relationship between codes assigned to documents and subgroups of cases defined by values of a numeric or categorical variable. When performing these kind of analyses, QDA Miner displays a contingency table containing either codes frequencies, code occurrences (present or absent), or either absolute or relative importance (in number or percentage of words). This dialog box also gives you access to several graphical and statistical tools to help you assess or visualize the strength of the relationship between code usage and the selected variable.



/	Bradley	Buchannan	Bush	Forbes	Gore	McCain	Student's F	P value
Local Economy		3	2	3		1	0,976	,492
Ethic	1				1		3,338	,074
Globalization		3	1	1	1		0,692	,646
Poverty		1			2		0,862	,550
Novelty	1	9	1	1	1	7	0,656	,668
Tradition	1			1		2	1,923	,209
Environment		5	2			1	0,710	,635

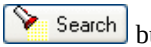
SEARCH IN - This option allows you to specify on which document variables the analysis will be performed. If the current project contains more than one document variable, you will have a choice of selecting either one of them or a combination of them. By default, all document variables are selected. To restrict the analysis to a few of them, click on the down arrow key at the right of the list box. A list of all available document variables will appear. Select all variables on which you want the search to be performed.

CODES - This option allows you to select which codes will be analyzed. By default, the analysis is performed on all codes in the codebook. You can select specific codes either from a drop-down checklist by clicking on the arrow button or from a tree representation of the codebook by clicking on the  button.

TABULATE WITH - This option list all numeric and categorical variables available in the project. Selecting any variable name will display a contingency table so that you can assess the relationship between this variable and the codes assigned to text segments.

To store the search expression and options, click on the  button and specify the name under which those query options will be saved.

To retrieve a previously saved query, click on the  button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

Once these options are set properly or when they are modified, click on the 

button to perform the analysis.

The remaining options are used to specify the information that should be displayed in the table.

COUNT - This option allows you to select which indicator will be used to quantify the importance of codes. Four indicators are available:

- Code Occurrence (case)
- Code Frequency.
- Word Count
- Percentage of Words

Selecting **Code Occurrence** allows you to compare the number of cases in which at least one instance of this code appears. If a code is used more than once in the same case, it will be counted only once. To count the total number of times a code has been used, select **Code Frequency**. Selecting **Word Count** will retrieve the total number of words that have been assigned to a specific code, while the **Percentage of Words** option divide this value by the total number of words found in those documents.

DISPLAY - The DISPLAY list box allows you to specify the information displayed in the table. The following options are available:

- Count
- Row Percent
- Column Percent
- Total Percent

When the COUNT option is set to **Code occurrence**, an additional statistic become available:

- Category percent (percentage of cases or individuals in this subgroup)

STATISTIC - This list box allows you to choose among 11 association measures to assess the relationship between the numeric or categorical variable and the codes.

Nominal level statistics

- Chi-square
- Student's F

Ordinal or internal level statistics

- Tau-a
- Tau-b
- Tau-c
- Somers' D (symmetric)

- Somers' Dxy (asymmetric)
- Somers' Dyx (asymmetric)
- Gamma
- Spearman's Rho
- Pearson's R

PROBABILITY - The probability option allows you to select whether the probability value should be calculated using a **1-tailed** or **2-tailed** test. Probabilities of Chi-square, Likelihood ratio, and Student's F are always calculated using a 2-tailed test.

After a first crosstabulation table has been obtained, additional buttons located on top of this dialog box will become available. These buttons are used to create barcharts or line charts representations of data in selected rows, perform correspondence analyses, or create heatmaps. You will find below short descriptions of the steps needed to access these features. A more detailed description of all these features will be presented later.

Creating Barcharts or Line Charts

- Barcharts or line charts are useful for visually comparing the distribution of specific codes among values of an independent variable such as subgroup of individuals (male vs. female) or time periods. To produce these types of charts:
- Set the TABULATE WITH, COUNT and DISPLAY options so that the information to visualize is displayed in the table.
- Using the mouse, select the rows that you would like to display. Multiple but separate rows can be selected by clicking while holding the **CTRL** key down.
- Click on the  button.

For more information, see **Barchart and Line Chart** on page 108.

Creating Heatmaps with Clustering of Rows and Columns

A heatmap plot is a graphic representation of crosstab tables where cell frequencies are represented by different color brightnesses or tones. When combined with clustering of rows and/or columns, this exploratory tools allows you to identify functional relationships between specific codes and subgroups defined by values of the independent variable. To create a heatmap:

- Set the TABULATE WITH, COUNT and DISPLAY options so that the information to visualize is displayed in the table.
- Click on the  button to access the heatmap dialog box.

For more information on heatmaps, see **Heatmap Plot** on page 111.

Performing Correspondence Analysis

Correspondence analysis is an exploratory technique that provides a graphic overview of relationships in large crosstabulation tables of frequency. To perform a correspondence analysis:

- Set the TABULATE WITH, COUNT and DISPLAY options so that the information to visualize is displayed in the table.
- Click on the  button to access the correspondence analysis dialog box.

For more information on correspondence analysis, see **Correspondence Analysis** on page 116.

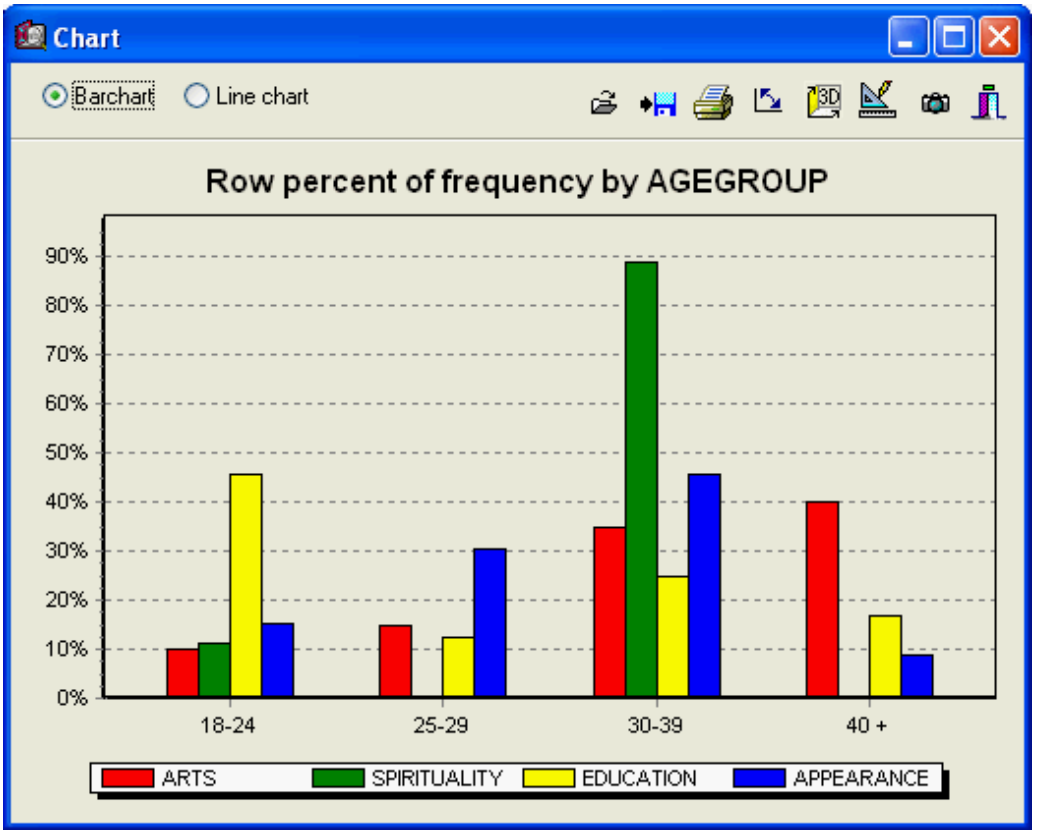
Printing and Exporting a Crosstabulation Table

The crosstabulation table may be saved to disk in Excel, MS Word, plain ASCII, text delimited, HTML, or XML format by clicking on the  button.

You may also print this table by clicking on the  button.

Barchart or Line Chart

The chart window allows you to graphically examine the relationship between codes and values of an independent variable. The barchart should be used to display the distribution of various categories within subgroups as defined by a nominal independent variable, while the line chart may be preferred for the examination of the relationship between these categories and an ordinal or quantitative variable.



The following table provides a short description of the available buttons and controls:

Control	Description
	Click this button to retrieve a chart previously saved on disk.
	Click this button to save a chart on disk. Charts may be saved in BMP, JPG or PNG graphic file format or may be saved in a proprietary format (.WSX file extension) that can later be edited and customized using the Chart Editor.
	Clicking this button prints a copy of the displayed chart.

Clicking this button causes the values represented on the bottom axis to be switched with those represented by different lines or bars (legend).

Click this button to turn on/off the 3D perspective for the current chart.

This button allows the editing of various features of the chart such as the left and bottom axes, the chart and axes titles, the location of the legend, etc. (see **Barchart and Line Chart Options** below)

This button creates a copy of the chart to the clipboard. When this button is clicked, a shortcut menu appears allowing you to select whether the chart should be copied as a bitmap or as a metafile.

Pressing this button closes the **chart** dialog box and returns to QDA Miner's main screen.

Barchart and line chart options dialog box

The various options in this dialog box are used to customize the appearance of barcharts and line charts. The options available in this dialog box represent only a small portion of all settings available.

To further customize the chart, edit data points, value labels, or series order, click on the  button located on the right side of the dialog box.

LEFT AXIS

MINIMUM / MAXIMUM - QDA Miner automatically adjusts the vertical axis scale to fit the range of values plotted against it. To manually set these values, enter the desired minimum and maximum.

INCREMENT - Increasing or decreasing this value affects the distance between numbers as well as tick marks. Horizontal grid lines are also affected by changing this value.

HORIZONTAL GRID - This option turns horizontal grid lines on and off. Grid lines extend from each tick mark on an axis to the opposite side of the graph. To increase or decrease the number of grid lines or the distance between these lines, change the Increment value of the axis. A list box also allows a choice among five different line styles to draw grid lines.

LEGEND

LOCATION - This option positions the legend. Legends may be placed at the Top, Left, Right and Bottom side of the chart.

FROM TOP - When the legend is displayed on the left or the right side of the chart, this option specifies the legend's top position in percentage of total chart height.

FROM LEFT - When the legend is displayed on the top or the bottom chart, this option specifies the legend's top position in percent of total chart height.

TITLES

Proper titles and axis labels are of utmost importance when describing the information displayed in a chart. By default, QDA Miner uses variable names and labels as well as other predefined settings to provide descriptions.

The title page allows you to modify the top title, as well as the labels on the left, bottom and right axes. To edit the title, select the proper radio button. Enter several lines of text for each title by pressing the **Enter** key at the end of a line before entering the next line.

The **Font** button on the right side of the edit box changes the font size or style of the related title.

3D VIEW

ORTHOGONAL - Turning this option off disables the free elevation and rotation of the 3D chart.

ZOOM - This option zooms in and out of the whole chart. Expressed as a percentage, this value, when increased positively, will bring the chart towards the viewer, increasing the overall chart size as the Zoom value increases.

3D PERCENT - The 3D Percent property indicates the size ratio between chart dimensions and chart depth by specifying a percent number from 1 to 100.

PERSPECTIVE - Use this property with Orthogonal unchecked to change the 3D perspective of the Chart. Larger values add more depth perspective.

BAR SHADOW - Enabling this option dark-shades the sides of 3D bars. Turning it off will paint the sides of the bar in the same color as the front of the bar.

BAR WIDTH - This option determines the percentage of total bar width used. Setting this value to 100 joins the bars.

BAR DEPTH - Use this property to limit the depth that each bar series uses. By default, bars will take up the part proportional to the number of bar series in the chart so that the back of a bar will join the front of the bar immediately behind it. To insert a gap between a series of bars, decrease this value.

Heatmap Plot

Heatmap plots are graphic representations of crosstab tables where relative frequencies are represented by different color brightnesses or tones and on which a clustering algorithm is applied to reorder rows and/or columns. This enter of plot is commonly used in biomedical research to identify gene expressions. When used in qualitative analysis, this exploratory data analysis tool facilitates the identification of functional relationships between related codes and group of values of an independent variable by allowing the perception of cells clumps of relatively high or low frequencies or of outlier values.

The heatmap plot implemented in QDA Miner allows you to graphically examine the relationship between codes (rows) and values of an independent variable (columns). While the clustering available through the CODING CO-OCCURRENCES command (see **Clustering Cases** on page 90) is performed on individual documents, the clustering analyses used in the heatmap plot are performed directly on the crosstabulation tables. As a consequence, the similarity index computed for two codes and used for clustering do not represent their co-occurrences within cases but measures the similarity of their distribution among the various groups of the independent variable. Likewise, two subgroups defined by values on the independent variable will be considered near to each other if the distributions of codes in those two groups are similar.

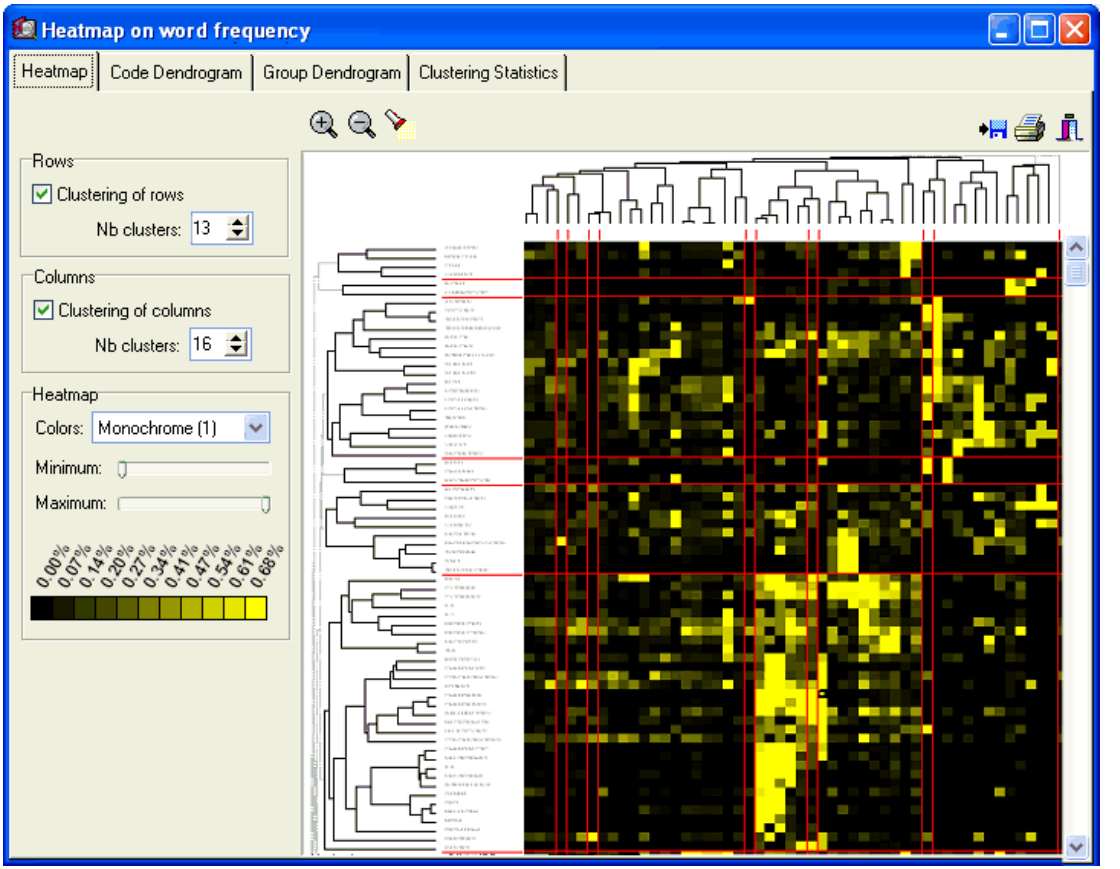
The heatmap plot can be performed on any of the available statistics (Code Occurrence, Code Frequency, Word Count, or Percentage of Words). For all these types of analyses, the observed frequency is transformed into a percentage by dividing the observed value by an appropriate statistic calculated for each subgroup of cases.

To create a heatmap:

- Select the CODING BY VARIABLES command from the ANALYSIS menu.
- Select the codes you want to plot along with the documents to which you wish to restrict the analysis.
- Set the TABULATE WITH option to the desired independent variable.
- Set the COUNT option to either **Code Occurrence**, **Code Frequency**, **Word Count** or **% of Words**.
- Click on the  button to construct the crosstab matrix.
- Click on the  button.

Heatmap Page

The main section on the right of this page, as shown in the screen shot below, displays the heatmap grid representing the relative frequencies of each cell (row and column intersection) using different brightnesses or color tones. Optional dendrograms are displayed at the top and on the left margin of this grid. The size of these dendrograms may be adjusted by moving the mouse cursor over the bottom edge (upper dendrogram) or the right edge (dendrogram on the left) of the dendrogram and dragging its limit to the desired size.



The font size used to display the row and column values may also be adjusted by clicking on the  or  buttons located on the top toolbar. The size of cells and the distance between dendrogram leaves are automatically adjusted to the new font size.

ROWS OPTIONS GROUP

CLUSTERING OF ROWS - Enabling this option reorders the rows based on the result of a cluster analysis and displays a dendrogram to the left of the heatmap plot. Codes that are distributed across the various subgroups in a similar way will tend to be grouped under the same cluster.

SORT BY - When code clustering is disabled, rows may be sorted in alphabetical order, in descending order of frequency or displayed in their original codebook order.

NO. CLUSTERS - This option allows you to set how many clusters the clustering solution should have. When two clusters or more are selected, horizontal red lines are drawn in the heatmap to delineate these clusters.

COLUMNS OPTIONS GROUP

CLUSTERING OF COLUMNS - Selecting this option reorders the columns based on the result of a cluster analysis and display a dendrogram at the top of the heatmap. Columns with similar distribution of codes will tend to be grouped under the same cluster. When this clustering option is disabled, the columns are presented in their ascending order of values. Disabling the clustering of columns is especially useful to preserve the ordinal nature of the values. For example, when looking for a relationship between some codes and publication years, then disabling the clustering of columns will automatically sort those columns in chronological orders allowing the identification of temporal trends.

NO. CLUSTERS - This option allows you to set how many clusters the clustering solution should contain. When two clusters or more are selected, vertical red lines are drawn in the heatmap to delineate these clusters.

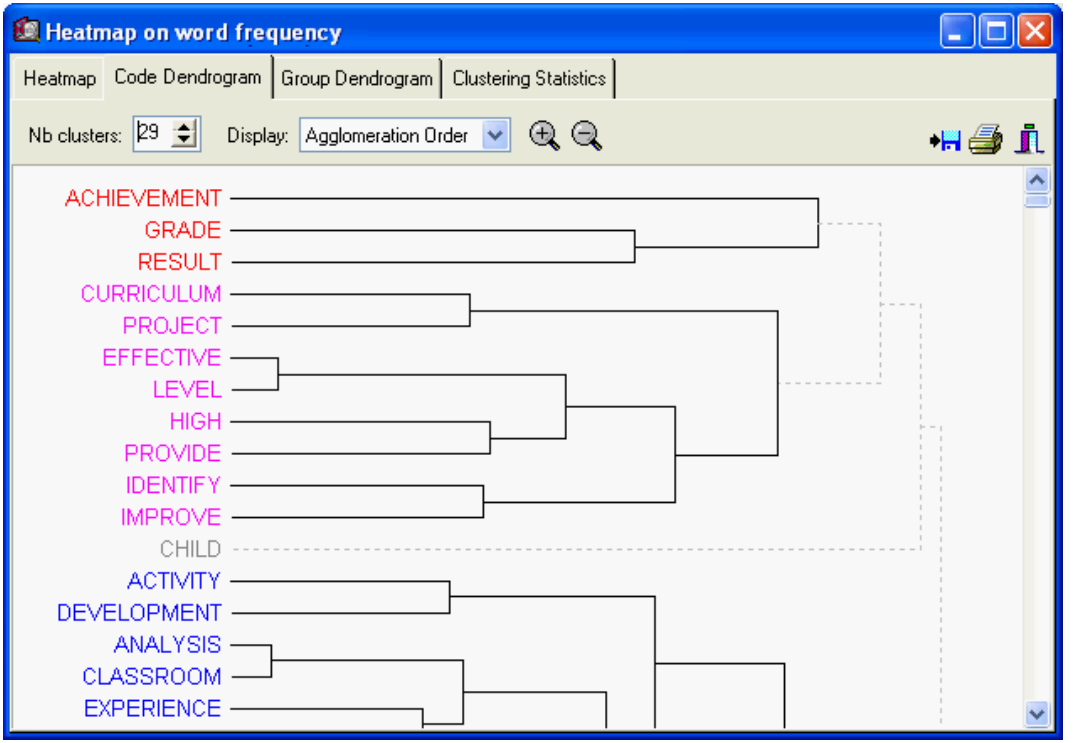
HEATMAP OPTIONS GROUP

COLORS - The colors list box allows the selection of various color schemes to represent differences in percentages. Monochrome schemes will express differences using levels of brightness for a single color where black always represents the lower limit of the selected range. Multicolor spectrums may be used to represent greater nuances in the selected range of percentages.

MINIMUM / MAXIMUM - The minimum and maximum slide bars allows you to set the range of values that will be used to display variations of color tones or brightness. By default, the minimum value is set to zero while the maximum value is set to the highest observed percentage. To increase the contrast at a specific location of this range, minimum and maximum limits may be adjusted. All cells with values lower than the minimum will be represented with the color located at the left end of the selected color spectrum, while cells with percentages higher than the maximum limit will be displayed with the color locate at the right end of this spectrum. Reducing the range of values increases the contrast in a specific region of percentage values.

Code and Group Dendrogram Pages

The second and third pages of the heatmap dialog box provide a more detailed examination of the clustering of codes (**Code Dendrogram**) and of all values on the independent variable (**Group Dendrogram**). QDA Miner uses an average-linkage hierarchical clustering method to create clusters from a similarity matrix. The result is presented in the form of a dendrogram (see below), also know as a tree graph. In this type of graph, the vertical axis is made up of the items and the horizontal axis represents the clusters formed at each step in the clustering procedure. Items that tend to be distributed similarly against the other variables appear together and are combined at an early stage, while those that have dissimilar distributions tend to be combined at the end of the agglomeration process.



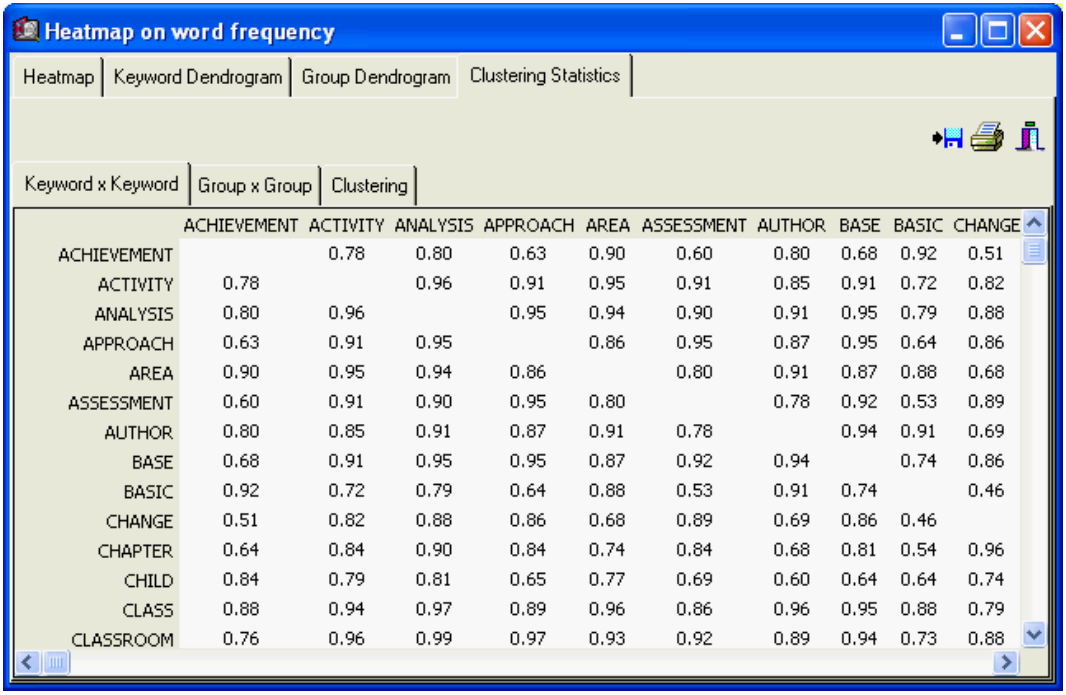
NO. CLUSTERS - This option sets how many clusters the clustering solution should have. Different colors are used in the dendrogram to indicate membership of specific items in different clusters.

DISPLAY - This option sets whether the vertical lines of the dendrogram represents the agglomeration schedule or the similarity indices.

At times, you will want to see some areas of a dendrogram. The  and  buttons may be used to zoom in or out of the dendrogram.

Clustering Statistics Page

The fourth page of the heatmap dialog box allows the close examination of matrices of similarities among codes or among groups as well as the statistics associated with the agglomeration processes of both the codes and the groups.



	ACHIEVEMENT	ACTIVITY	ANALYSIS	APPROACH	AREA	ASSESSMENT	AUTHOR	BASE	BASIC	CHANGE
ACHIEVEMENT		0.78	0.80	0.63	0.90	0.60	0.80	0.68	0.92	0.51
ACTIVITY	0.78		0.96	0.91	0.95	0.91	0.85	0.91	0.72	0.82
ANALYSIS	0.80	0.96		0.95	0.94	0.90	0.91	0.95	0.79	0.88
APPROACH	0.63	0.91	0.95		0.86	0.95	0.87	0.95	0.64	0.86
AREA	0.90	0.95	0.94	0.86		0.80	0.91	0.87	0.88	0.68
ASSESSMENT	0.60	0.91	0.90	0.95	0.80		0.78	0.92	0.53	0.89
AUTHOR	0.80	0.85	0.91	0.87	0.91	0.78		0.94	0.91	0.69
BASE	0.68	0.91	0.95	0.95	0.87	0.92	0.94		0.74	0.86
BASIC	0.92	0.72	0.79	0.64	0.88	0.53	0.91	0.74		0.46
CHANGE	0.51	0.82	0.88	0.86	0.68	0.89	0.69	0.86	0.46	
CHAPTER	0.64	0.84	0.90	0.84	0.74	0.84	0.68	0.81	0.54	0.96
CHILD	0.84	0.79	0.81	0.65	0.77	0.69	0.60	0.64	0.64	0.74
CLASS	0.88	0.94	0.97	0.89	0.96	0.86	0.96	0.95	0.88	0.79
CLASSROOM	0.76	0.96	0.99	0.97	0.93	0.92	0.89	0.94	0.73	0.88

Correspondence Analysis

Correspondence analysis is a descriptive and exploratory technique designed to analyze relationships among entries in large frequency crosstabulation tables. Its objective is to represent the relationship among all entries in the table using a low-dimensional Euclidean space such that the locations of the row and column points are consistent with their associations in the table. The correspondence analysis procedure implemented in QDA Miner allows you to graphically examine the relationship between assigned codes and subgroups of an independent variable. The results are presented using a two- or three-dimensional map. Correspondence analysis statistics are also provided to assess the quality of the solution. QDA Miner currently restricts the extraction to the first three axes. This ensures that the results remain easy to interpret. This limitation may however be removed in future versions upon request from users.

The first two pages of the dialog box provides graphical views of correspondence maps. When the number codes or the number of subgroups is less than four, only the **2D Map** page can be accessed. When the comparison is restricted to two groups of individuals, only one axis can be extracted. In this situation, the axis is plotted diagonally in the two-dimensional space. This has been done mainly for readability reason: plotting all the code labels on a single horizontal axis would have produced a cluttered list of codes that would have made the graph useless.

You will find immediately below a description of the various controls in the dialog box, followed by some guidelines for interpreting correspondence plots.

2D Map control

PLOT - When more than two axes have been extracted, this control allows the selection of all possible axis combinations that can be graphed on the two axes of the plot.

2D and 3D Map Controls

Codes This checkbox displays or hides the row points (i.e., code names).

Groups This checkbox displays or hides the column points (i.e., subgroup labels)

Clicking this button enables to zoom in on a plot. To zoom in on a area of the plot, hold the left mouse button down and drag the mouse down and to the right. A rectangle will appear around the selected area. Release the left mouse button to zoom in.

Clicking this button restores the original viewing area of the plot.

This button is used to create a copy of the chart on the clipboard. When this button is clicked, a shortcut menu appears, allowing you to choose whether the chart should be copied as a bitmap or as a metafile.

Clicking this button opens a dialog box to customize the appearance of the correspondence plots.

Clicking this button prints a copy of the displayed chart.

Clicking this button closes the correspondence analysis dialog box and returns to QDA Miner's main screen.

3D Plot controls

This button can be used to show or hide left, bottom and back walls.

Clicking this button draws anchor lines from the floor to the data point to better locate data points in all 3 dimensions.

Clicking this button changes the viewing angle of the chart. To rotate the chart, make sure this button is selected, click any area of the chart, hold the mouse button down and drag the mouse to apply the desired rotation.

Locating a data point on the depth dimension of a 3D plot can be very difficult, especially when the plot remain static. Often this plot must be rotated constantly on the various axes to obtain an accurate idea of where the data point is located on this third axis. Clicking this button rotates the plot automatically. To disable the automatic rotation, click a second time.

Interpreting Correspondence Analysis Results

Interpretation of correspondence analysis maps can be somewhat tricky and should be done with great care, especially when examining the relationship between row points and column points. Here are some basic rules that should help you interpret these maps:

Relationship among codes (row points)

- The more similar the distribution of a code among subgroups is to the total distribution of all codes within subgroups, the closer it will be to the origin. Codes that are plotted far from this point of origin have singular distributions.

- If two codes have similar distributions (or profiles) among subgroups of the independent variable (columns), their points in the correspondence analysis plot will be close together. For example, if the codes consist of artist names and the studied subgroups represent different age groups then, if the form of the distribution of two different artists among these age groups is similar, they will tend to appear near each other. Codes with different profiles will be plotted far from each other. Note: Two points may appear close to each other on a two- or three-axis solution, but may in fact be far apart when an additional dimension is taken into account.

Relationship among subgroups (column points)

- The more singular a profile of coding for a subgroup is, compared with the distribution of these codes for the entire sample, the farther this subgroup will be from the point of origin.

- If two subgroups of individuals have similar profiles of coding, they will be plotted near each other. Subgroups with different profiles will be plotted far from each other. Again, it is important to remember that two points may appear close to each other on a two- or three-axis solution, but may in fact be far apart when an additional dimension is taken into account.

Relationship between codes and subgroups (row and column points):

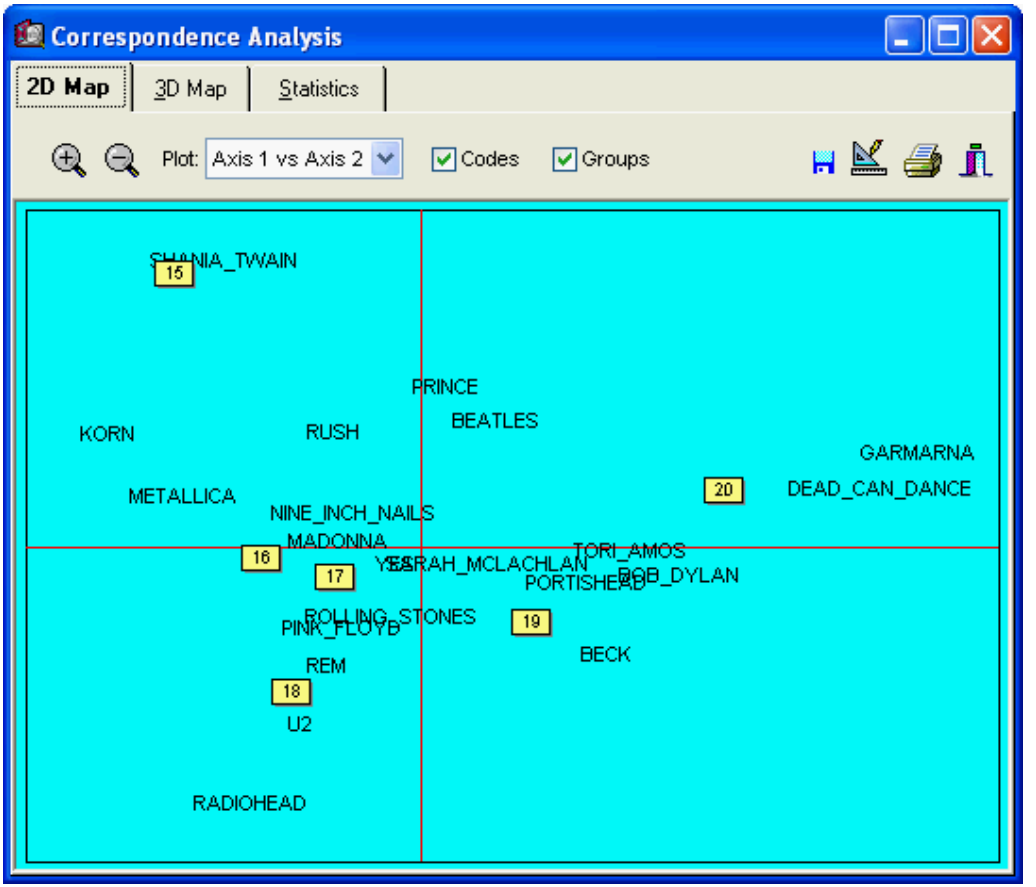
- Great caution should be taken when interpreting the distances between two sets of points (row and column points). The fact that the name of a subgroup is near a specific code should not necessarily be interpreted as an indication that they are closely related.

- While the distance between codes and subgroups has no interpretable meaning, the angle between this code point and a subgroup point from the origin is meaningful:

- An acute angle indicates that the two characteristics are correlated.

- An obtuse angle, near 180 degrees, indicates that the two characteristics are negatively correlated.

In the example below, REM, U2 and RadioHead could be viewed as related to 18-year-old listeners. However, despite the fact that both REM and U2 points are closer to the 18-year-old subgroup, Radiohead should be identified as more characteristic of those listeners since it is farther from the origin.



- Codes closely associated with two subgroups will be plotted at an angle from the origin that will lie between those two groups. In the above example, the rock groups Korn and Metallica seem to be characteristic of both 15 and 16-year-old listeners.

For a more complete description of this method, its calculation and applications, see Greenacre (1984).

Assessing Inter-coders Agreement

When document coding is performed manually by human coders, individual differences in interpretation of codes between different coders often occur no matter how explicit, unambiguous and precise the coding rules are. Even a single coder is often unable to apply the same coding rules systematically across time. One way to ensure the reliability of the application of coding rules is to ask different raters to code the same content or to ask a single rater to code the same document at different times. The comparison of codes is then used to uncover differences in interpretation, clarify ambiguous rules, identify ambiguity in the text, and ultimately quantify the final level of agreement obtained by these raters. Unfortunately, the application of inter-coders agreement procedures often involves some requirements or assumptions that are not always compatible with the qualitative data analysis processes. At least two compatibility problems can be identified:

1. The Codebook Problem - Computing inter-coders agreement requires that all coders share the same codebook. They should not be allowed to create their own codes or modify existing ones. An entirely inductive approach that would give each coder the possibility of developing his own codes during the analysis would result in a situation where the end result would very likely be incommensurable, each coder ending up with entirely different categories. In this situation, it would be impossible to establish the inter-raters agreement per se. Probably one of the only types of agreements that may be achieved in this case is by judging whether their final conclusions agree. While using a predefined unmodifiable codebook is the most obvious solution to this problem, a mixed solution may also be adopted. For example, the first coder may be given the opportunity to develop his own codebook or refine an existing one by adding categories or revising existing ones. Once the coding by the first coder is finished or once he is satisfied with the codebook state of development, this codebook can then be used by other coders to code existing documents. Another possibility would be to have a codebook that would contain a portion of unmodifiable codes but leave coders the possibility of modifying another portion of them or add new ones. In the latter case, only the fixed portion of the codebook could be submitted to an assessment of intercoders agreement.

2. The Segmentation Problem - Another problem is related to the common absence of predefined segments to be coded. The flexibility provided by many computer qualitative data analysis software to define the length and location of a segment to be coded creates a situation where the number of coded segments, their lengths and locations will very likely differ from one coder to the next. In this situation, a strict agreement definition that compares the assignment of the same codes with the same text segment would be too stringent to apply. The most common solution to this segmentation problem has been the prior segmentation of documents to be coded. However, an alternative solution is to somewhat relax the definition of the agreement.

QDA Miner allows to circumvent some of the constraints imposed by the inter-coders agreement procedures by providing different levels of agreement and by relaxing the criterion used to establish the overlap of segments..

But first, in order to assess inter-coder reliability, some prior settings and conditions are required:

- The project should be configured to support login by more than one user (for more information on how to define multiple users, see **Security & Multi-users Settings** on page 21).

- Each user with coding rights should then log into the project using his own user name and password in order to apply codes to documents.

- To ensure the independence of the coding process, a specific user should not be allowed to view code assignments made by other users. The QDA Miner **Multi-users Settings** page provides the option of hiding all codings made by other users.

- Each user has to code a common portion of the documents using the same codebook. The agreement can only calculated on documents that have been coded by more than one user. All other documents that have been coded by only one user will be ignored in the computation of the inter-coders agreement.

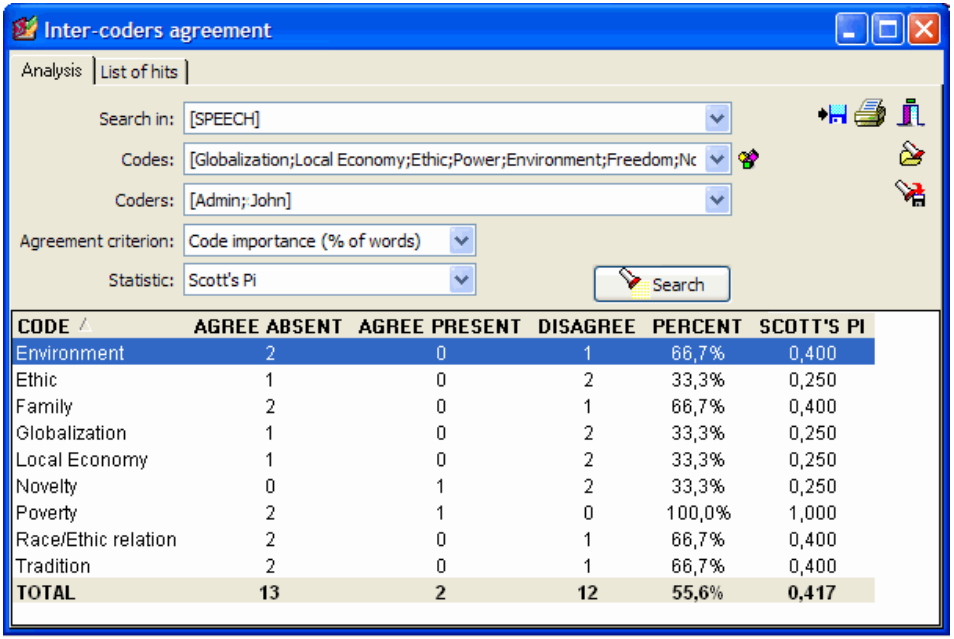
An alternative way to perform independent coding is to create multiple copies of an uncoded project, ask each coder to apply codes on their own copy of the project, and then combine all codings stored in different project files into a single master project using the project merging feature of QDA Miner (see **Merging Projects** on page 41). Each coder, however, MUST log on with their own user name

Only when all the above conditions have been met will it be possible to assess the agreement between coders.

Computing Inter-coders Agreement

In order to access the inter-coders agreement procedure, one should have the right to view codes created by other users. If a user do not have this right, only the administrator can remove this restriction by changing the multi-user settings.

To assess the inter-coders agreement, select the CODING AGREEMENT procedure from the ANALYSIS menu. A dialog box similar to the one below will appear:



CODE	AGREE ABSENT	AGREE PRESENT	DISAGREE	PERCENT	SCOTT'S PI
Environment	2	0	1	66,7%	0,400
Ethic	1	0	2	33,3%	0,250
Family	2	0	1	66,7%	0,400
Globalization	1	0	2	33,3%	0,250
Local Economy	1	0	2	33,3%	0,250
Novelty	0	1	2	33,3%	0,250
Poverty	2	1	0	100,0%	1,000
Race/Ethic relation	2	0	1	66,7%	0,400
Tradition	2	0	1	66,7%	0,400
TOTAL	**13**	**2**	**12**	**55,6%**	**0,417**

The following four options need to be set prior to an analysis:

SEARCH IN - This option allows you to specify on which document variables the analysis will be performed. If the current project contains more than one document variable, you will have a choice to select either one or more of them. By default, all document variables are selected. To restrict the analysis to a few of them, click on the down arrow key at the right of the list box. A list of all available document variables will appear. Select the variables on which you want the analysis to be performed.

CODES - This option allows you to select which codes will be analyzed. By default, the analysis is performed on all codes in the codebook. You can select specific codes either from a drop-down checklist by clicking on the arrow button or from a tree representation of the codebook by clicking on the  button.

CODERS - This option allows you to select the coders for which you want to assess the coding agreement. At least two coders must be selected. If more than two coders are selected, the agreement table will be calculated on all possible pairs of coders provided that they have coded the same documents. By default, this option is set to include all coders. To restrict the analysis to coded segments created by a few of them, click on the down arrow key at the right of the list box. A list of all coders will appear. Clear check marks beside the coders whom you want to exclude from the analysis.

AGREEMENT CRITERION - QDA Miner provides a way to assess four different levels of agreement. The Agreement Criterion option is used to select the level of agreement that will be tested by offering a choice of four criteria:

Code occurrence - This option checks whether coders agree on the presence of specific codes in a document, regardless of the number of times this code appears and of code location. The agreement is calculated on a dichotomous value that indicates whether the code is present or absent.

☞ Some coders may agree that something is present in a document without necessarily agreeing on what specifically are the manifestations of this "something". For example, a researcher may ask two coders to check numerous job interviews and ask them to decide whether there are some indications of anxiety from the candidate. They may very well agree in the identification of such anxiety despite their inability to agree on its specific manifestations. If a research hypothesis or an interpretation is related to the mere presence or absence of anxiety in various candidates, this level of agreement may well be sufficient. Coders do not necessarily have to agree on exactly where it occurs.

Code frequency - This option checks whether coders agree on the number of times specific codes appear in a document. This criterion does not take into account the specific location of the codes in the document. The agreement is calculated by comparing the observed frequency per document for each coder.

☞ If someone wants to examine how often a specific speaker uses religious metaphors in his speeches compared with other speakers, he may ask different coders to assign a code "Religious Metaphor" every time they encounter one in a speech. Even if they disagree somewhat on what represents a religious metaphor, they may nevertheless agree on how many of those metaphors they find in each speech. Since the research

question is related to the comparison of frequency of metaphors among speakers, then an agreement on the overall frequency per speech should be sufficient.

Code Importance - This option allows you to assess whether coders agree on the relative importance of a code in the document. This kind of agreement is achieved by comparing for a given code the percentage of words in the document that have been tagged as instances of this.

☞ If a researcher makes the hypothesis that when adolescents are asked to talk about a love relationship, male adolescents devote more time than females talking about the other person's physical characteristics and less time about their own needs or feelings. With this kind of hypothesis, the researcher needs to establish an agreement on the relative importance of these topics, not necessarily on the frequency or location. Again, coders may somewhat disagree on what specific segments represent instances where the adolescent is talking about their needs or feelings, but they may nevertheless come to a close agreement on the importance of each code.

Code Overlap - This last level of agreement is the most stringent criterion of agreement since it requires the coders to agree not only on the presence, frequency and spread of specific codes, but also on their location. When this option is selected, you need to provide a minimum overlap criterion expressed in percentage. By default, this criterion is set to 90%. If the codes assigned by both users overlap each other on 90% of their length, then these codes will be considered to be in agreement. If the percentage of overlap is lower than 90%, the program will consider them to disagree.

☞ Such a level of agreement would be interesting for several reasons. A researcher may be interested in indentifying specific manifestations of a phenomenon. Another important reason is that the examination of agreements and disagreements at this level is accessory to the achievement of high agreement on any other lower level. In other words, close examination of where agreements occur or do not occur should allow one to diagnose the source of disagreement and take corrective actions to establish a shared understanding of what each code means for every coder. In a sense, checking the agreement at this level may be used in a formative way to train coders. Yet, the final summative evaluation of agreement may still be performed on a less stringent level, provided that this lower level matches the research hypothesis.

The first three criteria use documents as the unit of analysis and calculated for each document a single numerical value per code and coder. This numerical value represents either the occurrence, the frequency of the specific code or the percentage of words in the document assigned to it. This Coding Analysis Features 80 measure is then used to establish the level of agreement. For example, when assessing agreement on code frequencies, if a user assigns a specific code to a document four times while another one uses this code only once for the same document, then the program will add 0.25 to the total number of agreements and 0.75 to the total number of disagreements. The absence of a code in the document is considered to be an agreement. Since the unit of analysis is the document and the numerical value for each code lies between 0 and 1, each document has an equal weight in the assessment of agreement, no matter how many codes are found in these documents.

When using code overlap as the criterion for assessing agreement, the unit of analysis becomes the assigned codes. For this reason, the total number of agreements and disagreements can be higher than for the other

three assessment methods. In this case, the absence of a code in the document is not considered to be an agreement.

STATISTIC - The simplest measure of agreement for nominal level variables is the proportion of concordant codings out of the total number of codings made. Unfortunately, this measure often yields spuriously high values because it does not take into account chance agreements that occur from guessing. Several adjustment techniques have been proposed in the literature to correct for the chance factor. Three of these are currently available in QDA Miner:

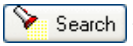
Free marginal adjustment assumes that all categories on a given scale have equal probability of being observed. It assumes that coder decisions were not influenced by information about the distribution of the codes. This coefficient is equivalent to the Bennett, Alpert and Goldstein's (1954) S coefficient, Jason and Vegelius's (1979) C Coefficient, and Brennan and Prediger's (1981) *k*n Index (Zwick, 1988).

Scott's pi adjustment (Scott, 1955) yields results that are similar to those obtained using Cohen's Kappa (Cohen, 1971). Like the Kappa, it does not assume that all categories have equal probability of being observed and considers that coders decisions are influenced by this information. However, unlike the Cohen Kappa, pi treats any difference among coders in this distribution as a source of disagreement.

Krippendorff's alpha (Krippendorff, 2004) is similar to Scott's pi but applies a correction for small sample sizes. The nominal alpha exceeds pi by (1-pi)/2n (where n is the number of units coded by the two coders) and asymptotically approximates pi as sample sizes become large.

To store the search expression and options, click on the  button and specify the name under which those query options will be saved.

To retrieve a previously saved query, click on the  button and select from the displayed list the name of the query you would like to retrieve (for more information, see **Saving and Retrieving Queries** on page 63).

Once the options has been set, click on the 

button to perform the analysis.

When codings by more than one user are available to assess the inter-coders agreement, two results tables will appear. The first table at the bottom of the **Analysis** page (see above) displays for each code various frequencies as well the level of agreement obtained. The table also provides information about the overall agreement. The content of this table depends on the level of agreement that has been chosen. You will find below the information such a table can contain:

COLUMN	DESCRIPTION
Code	Displays the name of the code.
Agree absent	Displays the number of times that users agreed on the absence of a code in a document. This column remains empty when assessing agreement using the code overlap criterion.
Agree present	Displays the number of times users agreed on the presence of a code in a document. When assessing agreement on code frequency or code importance, this

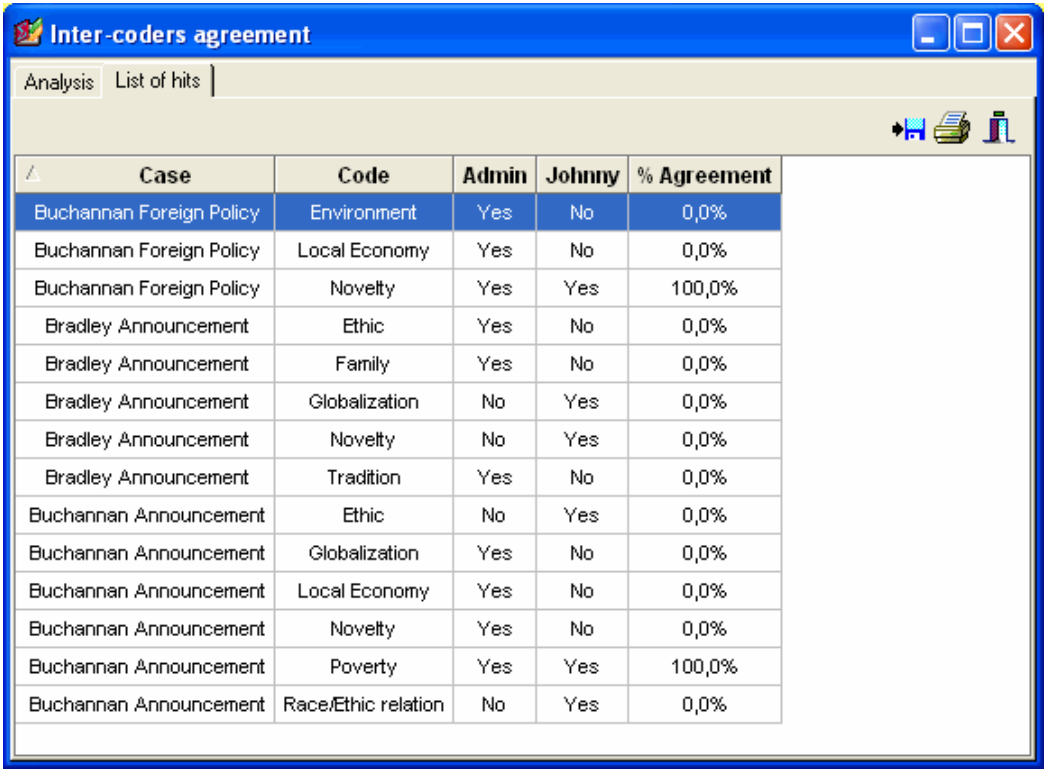
column displays the sum of all proportions of agreement calculated for each document.

Disagree — Displays the number of times users disagreed on the presence of a code in a document. When assessing agreement on code frequency or code importance, this column displays the sum of all proportions of disagreement calculated for each document.

Percent — This column reports the percentage of agreement observed between coders. This index is calculated by dividing the total number of agreements by the total number of comparisons possible (agreement + disagreement). This measure does not take into account chance agreements that occur from guessing. For this reason, the percentage of agreement often yields spuriously high values especially when assessing the less stringent levels of agreement.

The last column of the table displays the selected agreement statistic. This measure attempts to correct the observed level of agreement to take into account change agreements.

The second table located on the **List of Hits** page displays a list of all codings that were used for the assessment of inter-coders agreement.



Case	Code	Admin	Johnny	% Agreement
Buchannan Foreign Policy	Environment	Yes	No	0,0%
Buchannan Foreign Policy	Local Economy	Yes	No	0,0%
Buchannan Foreign Policy	Novelty	Yes	Yes	100,0%
Bradley Announcement	Ethic	Yes	No	0,0%
Bradley Announcement	Family	Yes	No	0,0%
Bradley Announcement	Globalization	No	Yes	0,0%
Bradley Announcement	Novelty	No	Yes	0,0%
Bradley Announcement	Tradition	Yes	No	0,0%
Buchannan Announcement	Ethic	No	Yes	0,0%
Buchannan Announcement	Globalization	Yes	No	0,0%
Buchannan Announcement	Local Economy	Yes	No	0,0%
Buchannan Announcement	Novelty	Yes	No	0,0%
Buchannan Announcement	Poverty	Yes	Yes	100,0%
Buchannan Announcement	Race/Ethic relation	No	Yes	0,0%

The table lists the case descriptor, the code name as well as the coders' values used for the calculation of agreement. When assessing code presence, code frequency or code overlap, this value will be either **Yes** (for present) or **No** (for absent). When assessing code importance, these columns will contain the total number of words assigned to this code.

Selecting an item in this table, either by clicking its row or by using keyboard cursor keys such as the **UP** or **DOWN** arrow keys automatically displays the corresponding case and document in the main window. This feature is especially useful for revising the codings made by different users and identify sources of disagreements.

The agreement table as well as the coding tables may be saved to disk in Excel, MS Word, plain ASCII, text delimited, HTML, or XML format by clicking on the corresponding  button.

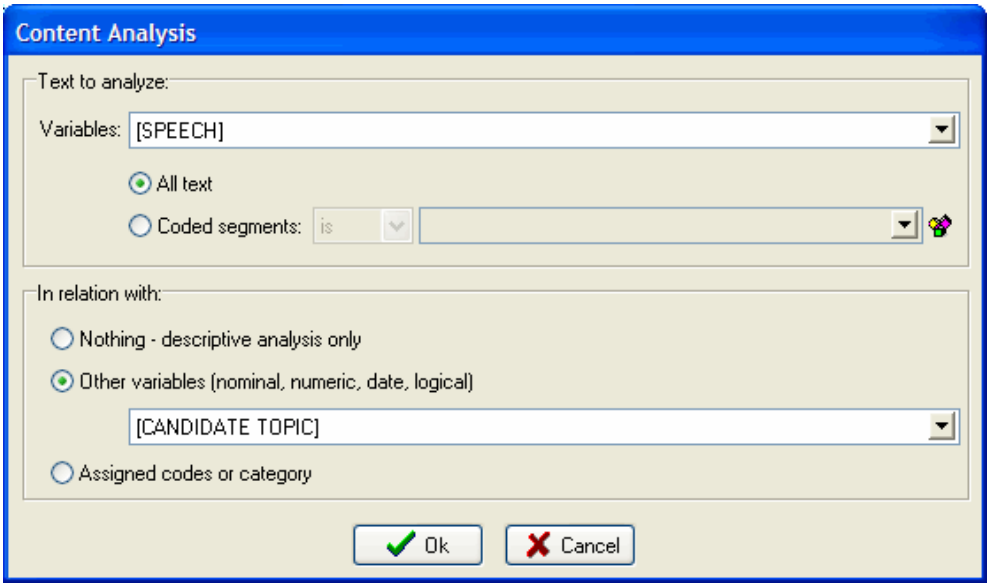
To print any table, click on the  button located on the tool bar above the corresponding table.

Content Analysis

WordStat is a quantitative content analysis add-on module that can be used to analyze words and phrases found in documents. When used with QDA Miner, WordStat can perform these analyses on entire documents or on selected code segments. It may also perform simple descriptive analysis or let you explore potential relationships between words, phrases or categories of words and other numeric or categorical variables.

There are many benefits to combining qualitative and quantitative content analysis techniques. For example, quantitative content analysis may be useful as an exploration tool prior to qualitative coding by allowing one to identify subtle differences in word usage between subgroups of individuals, or to quickly find the most common topics of phrases. Restricting the analysis to segments associated with specific codes may also be useful to identify potential words or phrases associated with those codes. You may then use the QDA Miner text retrieval tool to identify other segments to which this code may be assigned. Quantitative content analysis may also be useful after qualitative coding has been performed. For example, you may try to validate the conclusions drawn from manual coding by comparing those conclusions with the results obtained from a quantitative content analysis. Qualitative codings may also serve as the starting material to develop and validate a categorization dictionary that will allow automatic categorization of documents or cases.

To perform a quantitative content analysis on documents in a project, select the CONTENT ANALYSIS command from the ANALYSIS menu. The following dialog box will appear:



VARIABLES - This option allows you to specify on which document variables the analysis will be performed. If the current project contains more than one document variable, you will have a choiceof selecting either one or more of them. By default, all document variables are selected. To restrict the analysis to a few of them, click on the down arrow key at the right of the list box. A list of all available document variables will appear. Select all variables that you want to analyze.

By default, text analysis is performed on the entire document (**All Text** option). However, it is also possible to restrict the analysis to specific coded segments or exclude segments from the analysis by enabling the **Coded Segments** option. Set the list box immediately to the right of this option to **Is** if you want to analyze only the text found in the selected segments. Set it to **Is not** to exclude these text segments from the analysis. Finally, select the code segments to analyze or to exclude either from a drop-down checklist by clicking on the arrow button at the right end of the list box or from a tree representation of the codebook by clicking on the  button.

IN RELATION WITH - This set of options allows you to specify whether the quantitative content analysis should consider potential differences or relationship with one or more quantitative or categorical variables. If **Nothing** is selected, WordStat will allow descriptive as well as co-occurrence analysis on words or categories of words. Selecting the **Other Variables** option adds the possibility of examining in WordStat the relationship between these words or categories and the values of quantitative or categorical variables. The drop-down list box below this option is used to select the variables in relation to which the analysis will be performed. Selecting **Assigned Codes or Category** allows one to analyze the text assigned to codes and identify potential differences in words or categories of words associated with different codes. When this option is enabled, a temporary file is created containing all text segments associated with the selected codes along with a categorical variable to represent the code name. To add to this temporary file any remaining text segments not associated with any one of those codes, enable the **Include category for remaining text** option.

Once the options have been set, click on the **OK** button to call WordStat.

Statistical Analysis

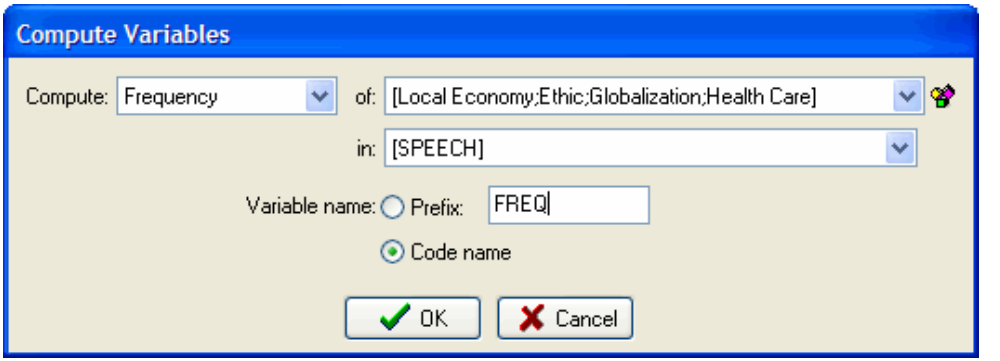
QDA Miner shares the same file structure and format as Simstat, an easy-to-use, yet powerful statistical program. When both applications are installed on the same computer, you can call Simstat from within QDA Miner and perform statistical analysis on any numerical or categorical variable in the project without having to export data to an external file, quit QDA Miner, or even close the active project.

Simstat offers a wide range of statistical analysis and graphics. It can also perform numeric and alphanumeric computation, transformation and recoding of variables, as well as advanced file management procedures such as data file merging, file aggregation, etc.

To run Simstat from QDA Miner, simply select the STATISTICAL ANALYSIS command from the ANALYSIS menu. While Simstat is running, QDA Miner will be disabled and it will not be possible to access any of its features. You will regain access to the program as soon as you quit Simstat.

Storing Code Statistics into Numeric Variables

Most output tables created by QDA Miner may be exported to disk in various file formats including Excel, MS Word, HTML, XML, and delimited ASCII files. Code statistics may also be appended to the existing project file, allowing you to use those new variables for further numerical processing or case filtering. To append code statistics to your current project, select the COMPUTE command from the VARIABLES menu. The following dialog box will appear:



COMPUTE - This option allows you to specify which statistic calculated on selected codes will be stored in the numeric variables. Four statistics are available:

- Code Occurrence (case)
- Code Frequency.
- Word Count
- Rate per 1000 words

OF - Select the codes on which to calculate the above statistic by clicking on the arrow button at the right end of the list box and by selecting these codes. The number of variables calculated will be equal to the number of codes selected .

IN - This option allows you to specify on which document variables the calculation will be performed. If the current project contains more than one document variable, you will have a choice to select either one or more of them. By default, all document variables are selected. To restrict the computation to a few of them, click on the down arrow key at the right of the list box. A list of all available document variables will appear. Select all the document variables that you want to analyze.

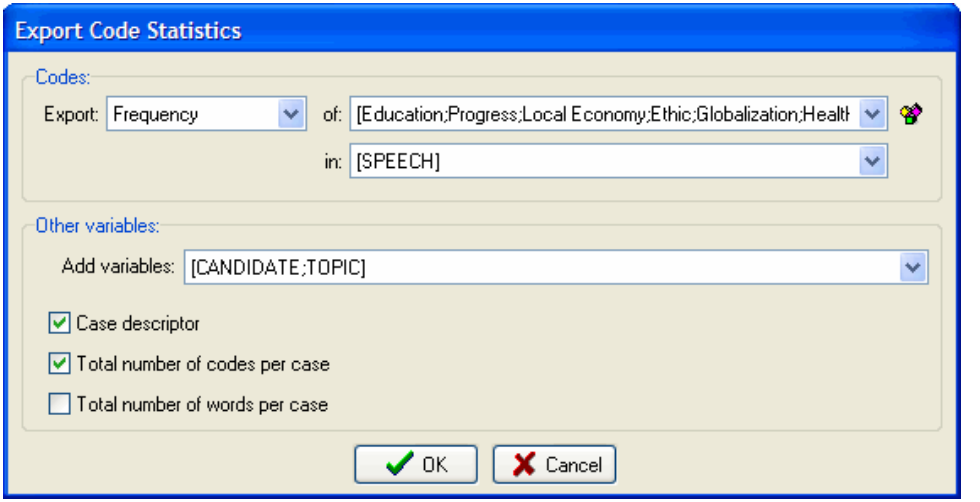
VARIABLE NAMES - This option sets what method should be used by QDA Miner to create new variable names. When set to CODE NAME, the program will attempt to use each code name as the name of a new variable. Illegal characters are automatically removed and long names are truncated to the first 10 characters. Duplicated variable names are distinguished by the substitution of numerical digits at the end of the name. When this option is set to PREFIX, variable names are created by adding successive numeric values to a user-defined prefix. For example, if the edit box at the right of the prefix option is set to "FREQ", the variable names will be FREQ1, FREQ2, FREQ3, etc.

When you click on the **OK** button, the program prompts you to confirm the calculation of the required number of variables. If variables with similar names exist in the project, the confirmation dialog box will display the number of existing variables that would be overwritten as well as the number of variables that need to be created by this command. Click on the **YES** button to confirm the creation and overwriting of variables. Click **NO** to abort the procedure.

To exit from the dialog box without performing the recoding, click on the **CANCEL** button.

Exporting Code Statistics

Various code statistics may be exported to disk in different file formats including Excel, HTML, XML and delimited ASCII files, allowing those data to be further analyzed using a statistical program. To export code statistics to disk, select the EXPORT | CODE STATISTICS command sequence from the PROJECT menu. The following dialog box will appear:



EXPORT - This option allows you to specify which statistic calculated on selected codes will be stored into numeric variables and exported. Four statistics are available:

- Code Occurrence (case)
- Code Frequency
- Word Count
- Percentage of Words

OF - Select the codes on which to calculate the above statistic by clicking on the arrow button at the right end of the list box and by selecting these codes. The number of variables calculated will be equal to the number of codes selected.

IN - This option allows you to specify on which document variables the calculation will be performed. If the current project contains more than one document variable, you will have a choice to select either one or more of them. By default, all document variables are selected. To restrict the computation to a few of them, click on the down arrow key at the right of the list box. A list of all available document variables will appear. Select all the document variables that you want to analyze.

ADD VARIABLES - This drop-down checklist box may optionally be used to add the values stored in one or more variables to the exported data file along with the code statistics.

CASE DESCRIPTOR - Enabling this option insert in the exported data file a string variable containing the case descriptor defined by the user. For information on how to create a case descriptor see .

TOTAL NUMBER OF CODES PER CASE - This options appends to each exported case, the total number of codes encountered in all selected documents.

TOTAL NUMBER OF WORDS PER CASE - This options appends to each exported case, the total number of words encountered in all selected documents.

Once the options have be set up, click on the **OK** button. A File Save dialog box will appear. Select the file format you want to create using the Save As Type drop-down list, enter a valid filename with the proper file extension and then click on the SAVE button.

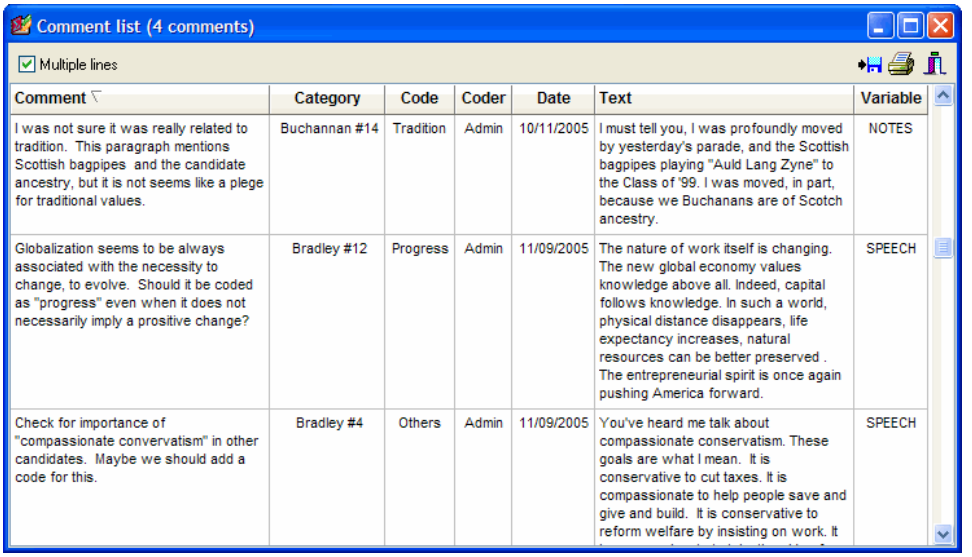
Exporting Co-Occurrence or Similarity Matrices

Matrices of co-occurrences or computed similarities may be exported to disk in Excel, MS Word, HTML, XML, and delimited ASCII files. To export such code co-occurence as well as code or case similarity matrix:

- Select the CODING CO-OCCURRENCES command from the ANALYSIS menu.
- Set the Co-occurrence options on the first page of the dialog box (see **Coding Co-Occurrences** on page 89 for information on those options)
- Move to the STATISTICS page.
- Click on the  button and select CO-OCCURRENCE MATRIX if you want to export the cross frequencies of codes, or SIMILARITY MATRIX to export the computed code or case similarity measures.
- Select the file format you want to create using the **Save As Type** drop-down list.
- Enter a valid filename with the proper file extension.
- Click on the **SAVE** button.

Retrieving a List of Comments

To review all comments attached to codings in a project, select the LIST COMMENTS command from the CODES menu. A dialog box similar to this one will appear:



Comment	Category	Code	Coder	Date	Text	Variable
I was not sure it was really related to tradition. This paragraph mentions Scottish bagpipes and the candidate ancestry, but it is not seems like a plege for traditional values.	Buchannan #14	Tradition	Admin	10/11/2005	I must tell you, I was profoundly moved by yesterday's parade, and the Scottish bagpipes playing "Auld Lang Zyne" to the Class of '99. I was moved, in part, because we Buchanans are of Scotch ancestry.	NOTES
Globalization seems to be always associated with the necessity to change, to evolve. Should it be coded as "progress" even when it does not necessarily imply a prositive change?	Bradley #12	Progress	Admin	11/09/2005	The nature of work itself is changing. The new global economy values knowledge above all. Indeed, capital follows knowledge. In such a world, physical distance disappears, life expectancy increases, natural resources can be better preserved . The entrepreneurial spirit is once again pushing America forward.	SPEECH
Check for importance of "compassionate convervatism" in other candidates. Maybe we should add a code for this.	Bradley #4	Others	Admin	11/09/2005	You've heard me talk about compassionate conservatism. These goals are what I mean. It is conservative to cut taxes. It is compassionate to help people save and give and build. It is conservative to reform welfare by insisting on work. It	SPEECH

Selecting an item in this table, either by clicking it or by using the keyboard cursor keys such as the UP or DOWN arrow keys, automatically displays the corresponding document in the main window and highlight the corresponding coded segment. This feature provides a way to examine the retrieved segment in its surrounding context.

To export retrieved comments to disk:

- Click on the  button. A Save File dialog box will appear.
- In the **Save as type** list box select the file format under which you would like to save the table. The following formats are supported: ASCII file (*.TXT); Tab delimited file (*.TAB); Comma delimited file (*.CSV); HTML file (*.HTM; *.HTML); XML file (*.XML); MS Word document (*.DOC); Excel spreadsheet file (*.XLS).
- Type a valid filename with the proper file extension.
- Click on the **SAVE** button.

To print the list of comments:

- Click on the  button.

References

Selected References on Qualitative Data Analysis

DEY, I (1993). Qualitative Data Analysis: A User-Friendly Guide for Social Scientists. Routledge.

KRIPPENDORFF, K. (2004). Content Analysis: An Introduction to Its Methodology. 2nd ed. Thousand Oaks, CA: Sage.

MILES, M. B. & HUBERMAN, M. (1994). Qualitative Data Analysis : An Expanded Sourcebook. Thousand Oaks, CA: Sage.

Inter-coders Agreement Statistics

BENNETT. E.M., ALPERT, R., & GOLDSTEIN, A.C. (1954). Communications through limited response questioning. Public Opinion Quarterly, 19, 303-308.

BRENNAN, R.L., & PREDIGER, D. (1981). Coefficient kappa: Some uses, misuses, and alternatives. Educational and Psychological Measurement, 41, 687-699.

COHEN, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological measurement, 20, 37-46.

JASON, S., & VEGELIUS, J. (1979). On generalizations of the G index and the phi coefficient to nominal scales. Multivariate Behavioral Research, 14, 255-269.

KRIPPENDORFF, K. (1970). Bivariate agreement coefficients for reliability of data. In E.F. Borgatta and G.W. Bohrnstedt (Eds.). Sociological methodology: 1970. San Francisco: Jossey-Bass.

SCOTT. W.A. (1955). Reliability of content analysis: The case of nominal scale coding. Public Opinion Quarterly, 19, 321-325.

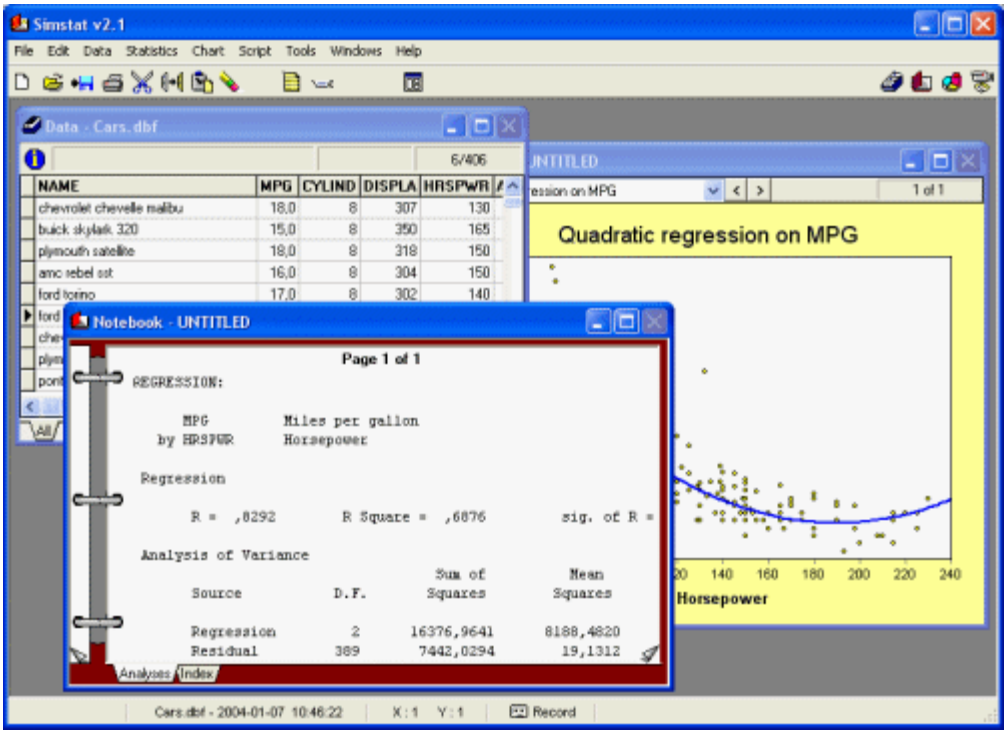
ZWICK. R. (1988). Another look at interrater agreement. Psychological Bulletin, 103 (3), 374-378.

Multivariate Analysis

GREENACRE, M. (1984). Theory and Applications of Correspondence Analysis. Academic Press. Orlando, Florida.

Information on Simstat

Simstat for Windows is a general statistical program. It offers a wide range of statistical and graphical tools, intuitive output management features as well as its own scripting language to automate statistical analysis, write small applications, interactive tutorials, etc.



For more information, visit Provalis Research web site at:

http://www.provalisresearch.com

Information on WordStat

WordStat is a text analysis module specifically designed to study textual information such as responses to open-ended questions, interviews, titles, journal articles, public speeches, electronic communications, etc. WordStat may be used for automatic categorization of text using a dictionary approach or various text mining procedures as well as for manual coding. WordStat can apply existing categorization dictionaries to a new text corpus. It also may be used in the development and validation of new categorization dictionaries. When used in conjunction with manual coding, this module can provide assistance for a more systematic application of coding rules, help uncover differences in word usage between subgroups of individuals, assist in the revision of existing coding using KWIC (Keyword-In-Context) tables, and assess the reliability of coding by the computation of inter-raters agreement statistics.

WordStat includes numerous exploratory data analysis and graphical tools that may be used to explore the relationship between the content of documents and information stored in categorical or numeric variables such as the gender or age of respondents, year of publication, etc. Relationships among words or categories as well as document similarity may be identified using hierarchical clustering and multidimensional scaling analysis. Correspondence analysis and heatmap plots may be used to explore relationships between keywords and different groups of individuals.

For more information, visit Provalis Research web site at:

http://www.provalisresearch.com

Technical Support

If you encounter any problems or have any commenst or suggestions for further improvement, please contact Provalis Research:

By Phone: 514-899-1672

By FAX: 514-899-1750

By Email: support@provalisresearch.com

Web site: www.provalisresearch.com

Mail: Provalis Research
2414 Bennett Avenue
Montreal, QC
H1V 3S4
Canada