ObsWin - Observational Data Collection and Analysis

Contents

Introduction

ObsWin is a software package for the collection and analysis of observational data. The software can be employed in a variety of research fields (e.g. Psychology, Ethology and any other field that lends itself to the recording and analysis of real time data).
Computerised data collection is now used extensively in a number of different settings, and with the availability of relatively cheap (and small) laptop and palmtop PC's, it is a quick, useful and viable alternative to pencil and paper.

Using the PC keyboard, up to 62 separate variables (e.g. behaviours) can be recorded in real time (either in live settings or from a videotape). Real time data collection is the best method of capturing the complexities of 'real life'. Each variable can be defined as either an event (something that happens quickly) or as a duration (something with an observable start and finish time): sometimes these are referred to as momentary events and duration events respectively. The coding scheme for the variables can be defined by the user before or after the observations. Templates of variable names can be created and applied in subsequent collection sessions, to avoid the user repeatedly defining variables on each subsequent observation session.

ObsWin has other advantages over paper and pencil schemes. There are options for coding mutually exclusive variables (i.e. groups of variables that cannot occur at the same time), an audible reminder option (to set the length of an observation session) and automatic saving of the datafile to safeguard against loss of data.
The resulting data are stored as an ASCII text file. The variables are recorded in the order that they occur, along with the onset and offset times for each. This allows different analyses of varying complexity to be conducted, as well as a number of graphical displays. On screen, the datafile appears as a table listing each occurrence of a variable, which can then be edited within ObsWin.

Analysis options include summary statistics (percentage occurrence. frequency etc.), a range of methods for sequential analysis (both time-based and event-based), inter-observer reliability and a number of graphical presentations/analyses.

The sample screen-shots and diagrams used throughout the manual are predominantly from the same datasets (EXAMPLE.SDS/EXAMPLE.OBS), shipped with the program. All graphs have been created using the ObsWin program itself.

Note This manual is designed to guide the user through the different features of ObsWin in a logical order. Menu paths are given, which show how to access procedures with the mouse, and screen shots from the working program are used for illustration purposes. It is assumed that the reader is already familiar with the Windows™ graphical environment.

Various statistical analyses are available (e.g. to determine the sequential order of variables) and are discussed briefly - for more detail the user should refer to the references given in the appendix. The guide is not intended to explain *why* data should be collected and analyzed in any specific way, but rather *how* to collect and analyze sequential observational data.

---

1. Starting Out

a) System Requirements

The minimum requirements needed to run ObsWin32 (32-bit Windows 95/NT version) are Windows 95 or Windows NT 4.0 and 8 MB of RAM.

b) Installing

To install ObsWin, place the CD into the CD drive (usually the *D* drive), select **Add/Remove Programs** from the Control Panel and click **Install**.
The default installation will create C:\Program Files\ObsWin32 as the parent directory for ObsWin.

c) Starting the program

Select **ObsWin32** from the ObsWin menu items via the Start menu.

Selecting the **Quick Start Data Collection** menu option launches ObsWin and automatically invokes the data collection screen to begin collecting data using the last used settings (see section 2).

Selecting the **ObsWin Help** menu option opens the ObsWin help file - this contains much of the information contained within this manual and is also available from within the program to provide online and context sensitive help.

Selecting **ObsWin Conversion Utility** invokes a program that may be used to convert datafile types independently from the main program.

1.1 The Startup Screen

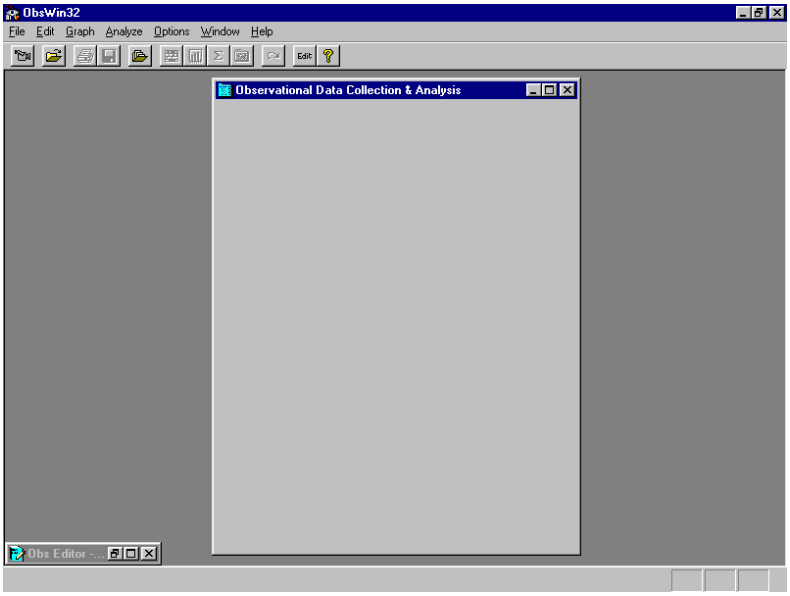


Fig. 1.1 - Start-up screen

When ObsWin is opened for the first time the datafile window will be blank. The user will see the menus, main toolbar, and Obs Editor icon (bottom left hand corner).

a) Menus

The pull down main menu and toolbar are at the top of the screen. ObsWin does not have its own programming or macro language, so the menus/toolbar buttons are the only way to run

procedures. For ease of location menu paths are given throughout this manual for every feature. They appear in bold print indicating the main menu item and corresponding sub-menu item:-

e.g. **File**
→ **Open**
→ **Datafile…**

Bold type without the arrow usually indicates a button on a dialogue box or toolbar (e.g. **Edit**), a caption or any other option.

b) Toolbar

Toolbars provide shortcuts to frequently used features (which are also accessible through the pull-down menus). Toolbars are used in many Windows applications (e.g. Excel, Word).

Figure 1.1 shows the main menu and toolbar. Some buttons on the toolbar are inactive (grey) because no datafile has been opened:

Choosing

Options
→ **Show Toolbar**

from the main menu will remove any toolbars from the display. This will remain in effect throughout subsequent sessions until the menu item is selected again (a tick will appear beside **Show Toolbar)**.

1.2 Data Output (the Obs Editor)

Obs Editor - OBS.LIS

The Obs Editor is like the SPSS Output window, or Minitab 'session window'. Most of the graphical and analysis options in ObsWin produce results or data that need to be output somewhere, and the Editor is used to store, view and scroll through this output. Files in the Editor are known as *listing files*: by default all output is saved to a listing file called OBS.LIS that resides in the default ObsWin directory.

Menu options shown in this section refer to the menus associated with the Obs Editor.

Clicking on the 

button on the main toolbar, pressing the **F4** key, clicking the Obs Editor icon or choosing

Edit
→ **Obs Editor**

from the main menu or pressing **F4** will invoke the Obs Editor and replace the main toolbar with the menu and toolbar shown in figure 1.2 below (the three buttons on the right of the toolbar are inactive until there are analysis data in the listing file).

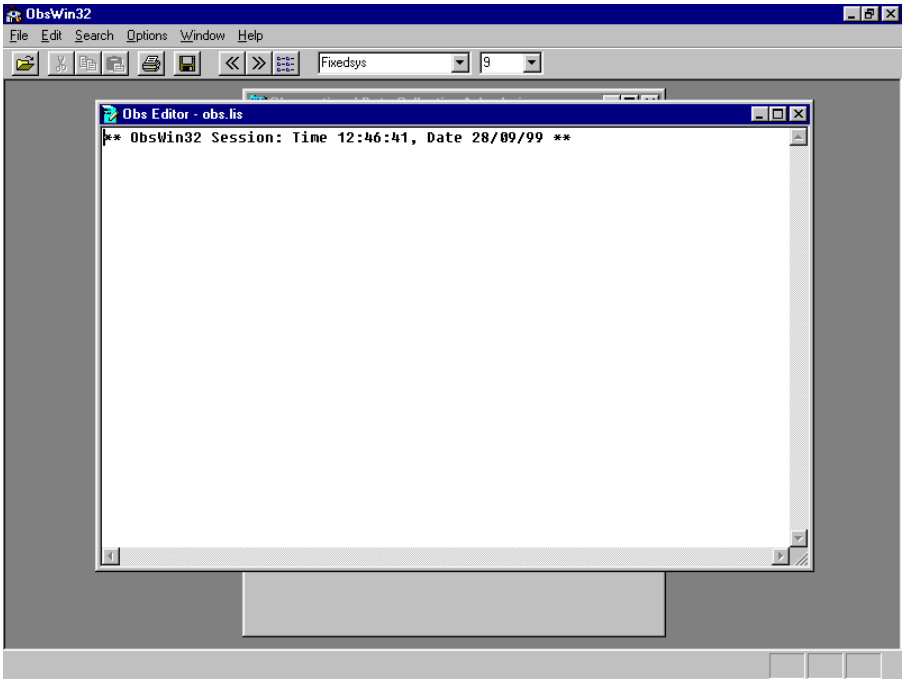


Fig. 1.2 - Obs Editor output window

a) Saving output (listing files)

Every time that you start ObsWin, (by default) a new empty OBS.LIS file is created in which to save current data. The date and time are shown at the top of the listing file, determined by the machine's internal clock (e.g. 'ObsWin Session: Time 4:02:44 PM, Date 5/2/99')
To keep the output you must rename the listing file to prevent it being overwritten in subsequent sessions. You are prompted to save the listing data when you finish an ObsWin session.

b) Printing output

Choose

File
→ **Print**

from the menu and the contents of the editor will be sent to the printer. There is no print confirmation option, so use

File
→ **Print Setup**

to check which printer is being used (especially important for users on a network).

c) Search menu

In many analysis packages, output from different procedures appears as one long piece of text. It is not always easy to spot the delimiters between them. The Obs Editor has the advantage of placing each output on a separate screen, and allowing the user to scroll through each of the entries saved (e.g. summary statistics, raw data from a lag analysis graph etc.).

When opened the editor automatically shows the data from the last analysis. Use the menu

Search
→ **View Next**
or → **View Previous**

or the toolbar buttons  or  to move back or forwards through the entries (see figure 1.2). Note that these options are only active when the editor contains more than one entry. The Editor also maintains a list of entries, accessed by choosing

Search
→ **View Listing Entry**

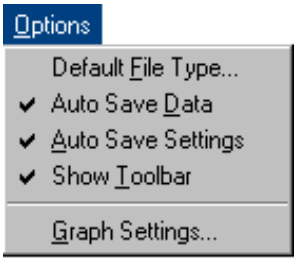
or choosing the  icon on the toolbar. The user can pick which entry to view (useful when the analysis has produced many outputs). The list is updated each time a new data entry is added/saved.

To leave the Obs Editor choose

File
→ **Exit Editor**

from the Obs Editor menu.

1.3 Environment Settings



The user can define certain ObsWin environment settings (e.g. observation settings, display colours for the graphs etc.). To save redefining these every time ObsWin is started, two options are available:

a) Auto-Saving Data

Checking the option

Options
→ **Auto-Save Data**

will ensure that the output generated from any of the analysis options will be automatically saved to the listing file (OBS.LIS) - this includes output from graphs as well as analysis data.

b) Auto-Saving Settings

Checking the option

Options
→ **Auto-Save Settings**

allows ObsWin to remember the current settings and re-load them.

2. Collecting observational data

2.1 Datafile Types

ObsWin will analyze a number of different file types, all but one are ASCII text files, and present variables that have an onset and an offset time, are ordered sequentially, and represent data that have been recorded continuously or by using a partial interval recording method.

a) The default datafile type used by ObsWin is the Timed Sequential datafile, one of four types of the Sequential Data Interchange Standard (*.SDS datafiles) developed by Bakeman and Quera (see references). This data format is widely used to represent data collected in continuous time and allows datafiles created using ObsWin to be analyzed in other packages and vice versa.
SDS datafiles may consist of numerous datasets. There may be more than one dataset separated by semi-colons, which are referred to as subject datasets, or more than one group of subject datasets which are referred to as sessions and separated by a forward slash. If you try and open one of these SDS datafiles then you are shown the number of subject/session datasets so that a specific dataset may be picked.

Note To analyze more then one of these datasets then you must create a fileset and add the datasets that you want to analyze to the fileset - see Using Filesets, section 3.4).

---

b) *.OBS datafiles are a very simple comma delimited text file format making them easy to edit or type using a text editor or word processor - each line has the variable name, its onset and offset separated by commas - any line that does not have two separating commas is ignored and so can be used for annotation/notes. If the last variable offset interval is not the last interval of the observation then this should be denoted by the variable END with the onset and offset interval being the last observation interval.

c) ***.**OBC (OBS compressed) datafiles are binary files that store the data in a similar way to OBS datafiles although taking up far less valuable disk space.

d) *.BEH datafiles are created by Observe (an observational data collection program written by Communitech International) and have a simple text file format similar to the OBS format, although some of the header information is used (i.e. the total time/number of intervals of the file). There are a few different versions of the *.BEH datafile and so when one is opened the first time the total time header must be identified - this configuration information is saved for subsequent loading of *.BEH datafiles.

e) *.ODB datafiles are created by CTS, the Tizard Centre's (University of Kent) continuous observational data collection program, written for the popular Psion Organizer. The datafiles consist of a number of lines of header information followed by a variable and the interval in which it was recorded. Header information determines which variables are events and which are durations so that duration offset times are simply determined by the next instance of that particular variable.

f) *.ODF datafiles are created by The Observer (version 3.0, Noldus Technology), a software package for the collection of different types of observational data. The datafiles consist of a number of lines of header information followed by the time interval and variable recorded at that interval. Configuration information (definitions for mutually exclusive sets/classes of variables and the variable definitions themselves) are stored in a separate configuration (*.CNF) file. ObsWin will only read in focal sampled (continuously recorded) datafiles created by version 3.0 of the Observer.

g) *.ABC datafiles are created by the Apple application ABC Data Pro (for iPones, iPads and iPod Touch devices – CBTAonline), software for the collection of different types of observational data. The datafiles consist of text files that the user emails to themselves after data collection and saving the data. The email text must be saved as a text file with the *ABC extension.

All datafiles have the following features:

- Continuous, real-time recording (or continuous partial interval recording) of variables.
- Variables are either events or durations with specific onset and offset times.
- Variables are recorded in the sequence in which they occur.

Choosing a Default File Type

Selecting

Options
→ **Default File Type**

from the main menu opens the **Default File Type** options box (figure 2.1) and allows the user to choose a default datafile type (i.e. *.SDS, *.OBS, *.OBC, *.BEH, *.ODB, *.ODF or *.ABC). This default file type will then be used for opening datafiles.

Note ObsWin datafiles may only be saved in *.SDS, *.OBS or *.OBC format.

2.2 New Observations

Click on the **New Observation** icon on the main toolbar  . Alternatively choose

File
→ **New Observation**

This will open the **Observation Settings** dialogue box (see figure 2.2), which allows the user to define a coding scheme. The user can specify variables as events or durations and give descriptive labels, as well as other settings shown below.

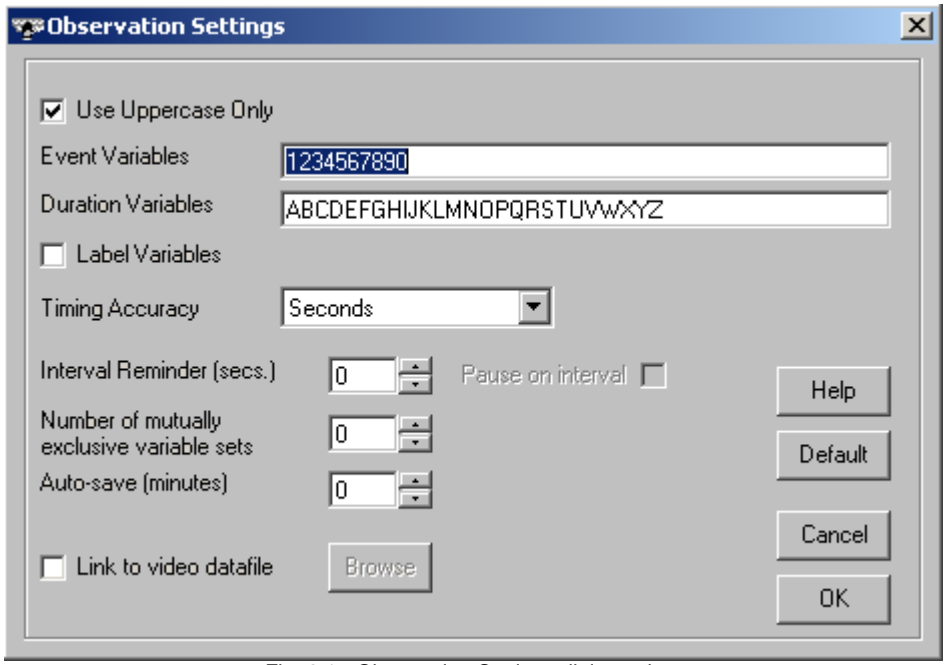


Fig. 2.1 - Observation Settings dialogue box

The advantage to defining the variables beforehand is that ObsWin displays a 'reduced' keyboard on the data collection screen, uncluttered by the keys that won't be used, and the labels are shown for reference (see figure 2.3).

2.3 Data Collection Settings

The **Observation Settings** dialogue box contains seven options. At least one variable must be defined to proceed.

Observational data are recorded using either the PC keyboard or with the mouse. By choosing **OK** on the observation settings, a keyboard image is displayed on the screen which can contain up to 62 separate keys (letters and numerals - see figure 2.2)**.** Each key is assigned to a different variable**.**

(a) Use uppercase only

By default the collection program is not case sensitive. i.e. only 36 variables can be defined (upper case variables only), and the use of the **Shift** key and **Caps Lock** are ignored. Pressing any valid key will record an upper case variable. Clearing the **Use Uppercase Only** box causes the program to read upper and lower case letters for different variables. This adds a further 26 variables to the possible range (number keys are not affected by the use of the shift key). The keyboard image used will then include extra lowercase keys if chosen. Obviously extra care against errors is needed during case sensitive keyboard recording.

There are 3 steps in defining variables:

- Decide which keys will represent which variables,
- Define each key as an event or a duration,
- Enter the letter/number once into the appropriate box

(b) Event (momentary) variables

Event variables are momentary variables that occur within a time interval (the offset time interval will be the same as the onset time interval). Variables that occur quickly should be defined as events (e.g. entering a room, dropping an object, throwing an object). One key press records the occurrence of an event, so these variables will form frequency records rather than duration measures.

Note By default, all the number keys are used as event variables.
All datafile types other than SDS allow multiple occurrences of the same variable within the same interval. If the datafile is saved as an SDS file, then any such occurrences are deleted during saving (although the user is given the option of either deleting such occurrences or re-saving as another file type).

(c) Duration variables

Durations variables have an onset and an offset time interval (e.g. walking around, eating, scratching). The duration measured is the number of time intervals in which the variable was occurring. Press the assigned key to indicate the start, and press again to indicate the finish (the button corresponding to the variable will remain depressed when the variable is ongoing - see figure 2.3 where the letters 'A' and 'S' are ongoing).

(d) Label variables

Checking this option allows the user to label variables prior to data collection. The user can enter meaningful names up to 10 characters long. Labels must be unique. Variables can also be labelled later by choosing

Edit
→ **Rename Variables**
→ **Edit Variables…**

from the main menu. Labels will appear on screen whilst recording (see figure 2.3).

When the observation settings have been chosen, click on the **OK** button. If the user has requested variable labels, the dialogue box for these will appear. After these are set the program will proceed to the data collection window (see figure 2.3), ready for the session.

(e) Timing accuracy

The timing accuracy for data collection may be set at seconds, tenths of seconds or hundredths of seconds.

Note The degree of accuracy used to collect data may be influenced by the ability (or otherwise) to establish inter-observer reliability!

(f) Interval reminder (secs.)

The user may set an interval reminder of *n* seconds (e.g. to indicate a new condition). Use the arrow buttons (increments of 1s) or enter a time in the box. When an interval has been set (any value greater than zero), either the computer will beep every *n* seconds during data collection or, if the 'Pause on interval' box is checked, the data collection timer will pause at the end of each interval. This pause function is most useful when linked to a digital video where the video is also paused allowing the user to consider how to code the interval just viewed (time sampling).

(g) Mutually exclusive variable sets

These should be used where a number of non-overlapping variables are being recorded (e.g. presence/absence of an individual). Only one variable in a mutually exclusive set can be occurring at any time.
Mutually exclusive variables are recorded in the same way as durations except that if one variable in the set is ongoing and another begins it is not necessary to turn off the first variable - this will happen automatically when the next key is pressed.
If the user has requested mutually exclusive variable sets the dialogue box for these will appear. This allows one to define each mutually exclusive set, and to choose colours to differentiate between them on screen.

Note Mutually exclusive variable sets must consist of variables that have been already declared as duration variables. Events cannot (by definition) be mutually exclusive. There may be a maximum of 7 sets - the default is to have no mutually exclusive sets.

(h) Auto-save

The **Auto-Save** feature is useful if data are being collected on a laptop or other battery powered PC. The data can be automatically saved every *n* minutes.
Use the arrow buttons (increments of 1 minute) or enter a time in the box and the computer will automatically save the data every *n* minutes during data collection.

Note Auto-saved data are saved to the file COLLECT.TMP located in the default ObsWin directory (with a normal installation this will be C:\OBSWIN), and will be deleted at the end of an ObsWin session (unless of course the session ends unexpectedly due to battery failure).

(h) Link to video datafile

Checking this option allows one to select (using the **Browse** button) a digital video file (MPEG, AVI, QT, MOV or VOB formats) that is linked to data collection. The video will run on screen with the virtual keyboard (see figure 2.3 below) and be controlled by the function keys F2 – F9 (see next section).

Note Your PC must be capable of watching DVDs in order to use the DVD file type to link to observation. Commercial DVD files are protected and will not be able to be played. It may be necessary to install the MediaPlayer software included on the CD

2.4 Data Collection

Figure 2.3 below is an example of the data collection window. It displays all the essential features of recording.

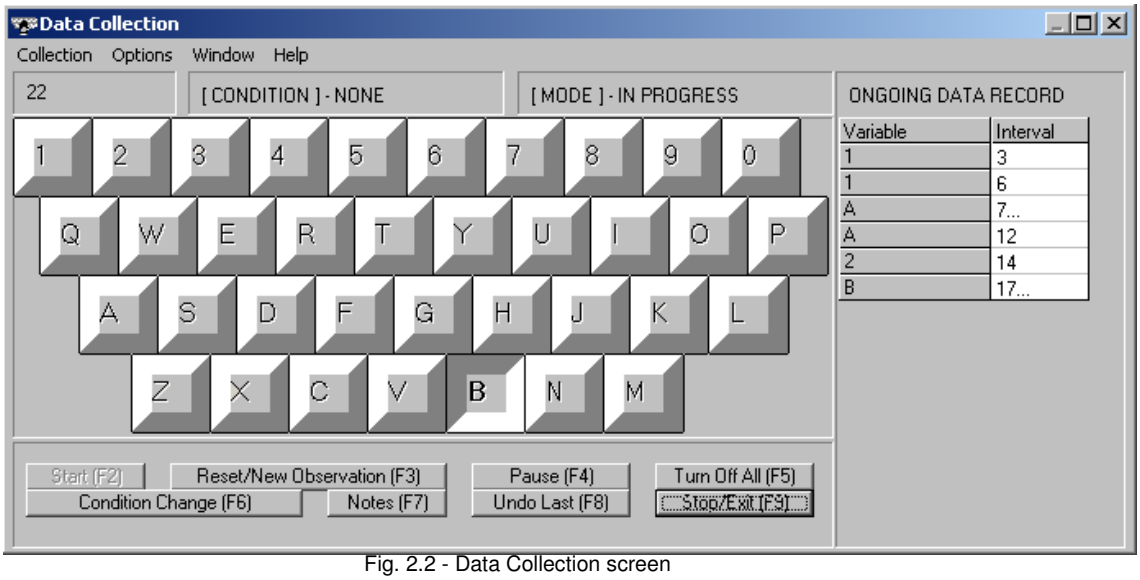


Fig. 2.2 - Data Collection screen

Menus

There are four pull down menus on the data collection screen. The **Collection** menu accesses the function keys shown at the bottom of the screen and detailed below, **Options** allows the user to set up a new observation session and/or edit existing observation settings.

Panels

There are three upper panels. These display the elapsed number of intervals of the observation session (22 intervals in the example shown in figure 2.3), whether any observation condition is currently in effect (see F6 function key below) and whether the observation session is in progress or has been paused (see F4 function key below).

The main panel displays the virtual keyboard. All keys selected to be used are displayed.

Note If you have chosen to label the variable keys then the labels can be displayed as floating text by hovering over the relevant key with the mouse.

The Ongoing Data Record panel to the right shows a record of the current observation session by simply displaying the key and the interval in which the key was pressed. Duration onsets are denoted by … following the onset. The example shows that the duration key B is currently depressed and that its onset occurred during interval 17 (e.g. '17…').

The bottom panel displays the function keys that can be used to control other aspects of the observation sessions.

Function Keys

F2 - Start

Initially the keyboard is inactive. To begin the recording press **F2** or click the **Start (F2)** box using the mouse. The timer at the top left will begin counting and the mode indicator will read 'In Progress' (see figure 2.3). Any key presses will now be recorded as variables and

stored in memory. The variables can be entered either directly from the keyboard, or by using a mouse to click on the screen image. If using a video file the video will start running.

F3 - Reset/New Observation

Pressing the **F3** key during data collection will automatically turn off any and all ongoing variables (their offset times are recorded at the time interval that **F3** was pressed), and the data saved as OBSn.NEW (where n is the next number in sequence if multiple datafiles are saved in this way).

Note If datafiles named OBSn.NEW are not renamed after their collection then they may be overwritten in subsequent sessions. There is a prompt to rename the datafiles when the session ends.

F4 - Pause

Pressing the **F4** key during data collection will stop the timer and pause the recording until it is pressed again - the timer will then continue from where it stopped. This is useful where the user is using videotape or recording from a video file rather than collecting data in live settings.

F5 - Turn Off All

Pressing the **F5** key during recording will turn off all ongoing variables (setting their offset times to the time interval that F5 was pressed). This is useful when an observation is temporarily obscured. The data recorded so far are not saved automatically (contrast this with **F3 Reset**).

F6 - Condition Change

Pressing the **F6** key during data collection will turn off any ongoing variables and open a text box. Type in a name/reference to a condition here (max. 5 characters). A condition can be any relevant change the user wishes to consider in later analyses (e.g. person left alone in room, lights switched on etc.). Press **Enter** to continue, and the name will appear in the condition indicator (top centre - see figure 2.3).

The onset of the condition will be the time interval when the F6 key was pressed and the offset interval will be the next time it is pressed (i.e. the start of another condition). The use of conditions is similar to the use of a mutually exclusive variable set except that condition variable names need not be determined in advance.

The condition will appear as a variable entry in the datafile at the appropriate time interval. Conditions may be marked on some of the graphical displays (e.g. occurrence graph).

This option allows ObsWin to analyze data by the presence or absence of a particular condition (see **Analysis Options** in section 6.9).

F7 - Notes

Pressing the **F7** key during data collection opens a text box at the top, in which notes can be made during the session. These notes are saved within the datafile with the time interval at which the F7 key was pressed. The notes can be marked on an occurrence graph as markers (see **Graphing Data** in section 5).

F8 - Undo Last

Pressing the **F8** key during data collection will undo the last key press - if this was an event key then the event variable is erased from memory, if it was a duration variable the onset or offset is erased, and the visual appearance of the key on the screen will change

appropriately. If the **F5** or **F6** keys turned off any variables then these will be switched back on.

Note Pressing the **F8** key again restores what was just undone i.e. whatever the F8 undo key has just undone will itself be undone. This guards against the possibility of the F8 key being hit in error.

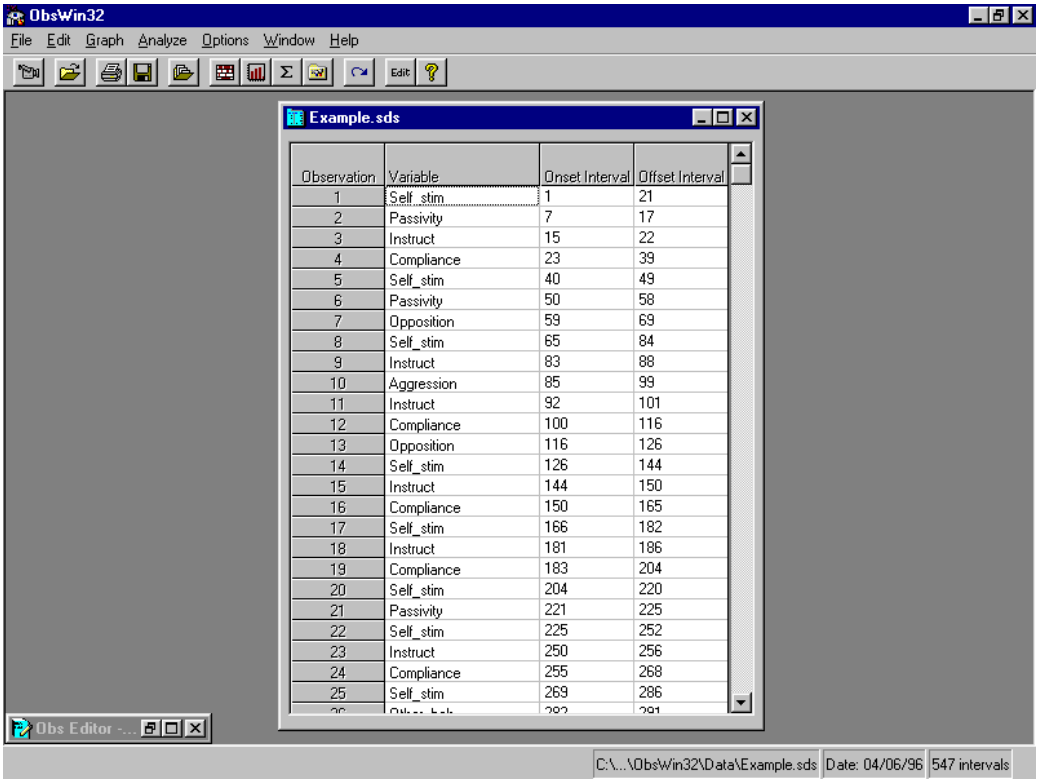
F9 - Stop

Pressing **F9** ends the timer, closes the **Data Collection** window (prompting to save any datafiles saved using **F3**) and displays the last datafile onscreen.

3. Manipulating datafiles (The File menu)

3.1 Opening/Viewing Datafiles

Once the user has pressed the **F9** key during data collection, the Data Collection window closes and the datafile is displayed in a spreadsheet-like grid as shown below in figure 3.1.



Observation	Variable	Onset Interval	Offset Interval
1	Self_stim	1	21
2	Passivity	7	17
3	Instruct	15	22
4	Compliance	23	39
5	Self_stim	40	49
6	Passivity	50	58
7	Opposition	59	69
8	Self_stim	65	84
9	Instruct	83	88
10	Aggression	85	99
11	Instruct	92	101
12	Compliance	100	116
13	Opposition	116	126
14	Self_stim	126	144
15	Instruct	144	150
16	Compliance	150	165
17	Self_stim	166	182
18	Instruct	181	186
19	Compliance	183	204
20	Self_stim	204	220
21	Passivity	221	225
22	Self_stim	225	252
23	Instruct	250	256
24	Compliance	255	268
25	Self_stim	269	286

Fig. 3.1 - Open datafile (EXAMPLE.SDS)

Alternatively, one may open an existing datafile by choosing

File
→ **Open**
→ **Datafile…**

on the main menu, pressing the **F2** key or clicking the  button on the main toolbar. Only one datafile can be opened at a time. If changes are made to the open datafile the user will be prompted to save the file on exit.

3.2 Creating Datafiles

A new datafile can be created from the main menu, without collecting data from an observation session, by choosing

File
→ **Create**
→ **Datafile…**

This allows the manual input of data which may be useful when there are existing data from an outside source that cannot easily be transformed into one of the compatible data formats (*.SDS, *.OBS, *.OBC, *.BEH, *.ODB [CTS] or *.ODF).

A new datafile template is created and requires the editing of cells to input data (see section 4 - **Editing Data**).

Any time intervals entered are checked for errors. If a variable is out of sequence there will be a warning displayed when any procedure is attempted.

3.3 Saving and Printing Datafiles

Choosing

File
→ **Save Datafile**

or clicking the  button on the main toolbar saves the datafile. If the datafile is new (UNTITLED) then the user will be prompted to specify a filename - this is equivalent to choosing

File
→ **Save Datafile As…**

By default files are saved to the C:\OBSWIN\DATA directory. Optional descriptive headers may be included for each datafile - check the **Add Header** box.

Choosing

File
→ **Print Datafile**

or clicking  on the main toolbar will send the datafile to the printer. There is no print confirmation option, so use

File
→ **Print Setup…**

to check which printer is being used and to change other print settings.

3.4 Using Filesets

Filesets are simply a collection of datafile names which can be analyzed together. This is not the same as combining or appending files (see section 4.10). Filesets save time by allowing the user to run analyses or create summary graphs over a group a files at the same time, rather than analyzing each datafile separately.

Choosing

File
→ **Open**
→ **Fileset**

opens a fileset.

Choosing

File
→ **Create**
→ **Fileset**

or clicking  on the main toolbar invokes the Create Fileset dialogue box (see figure 3.2) to create a fileset.

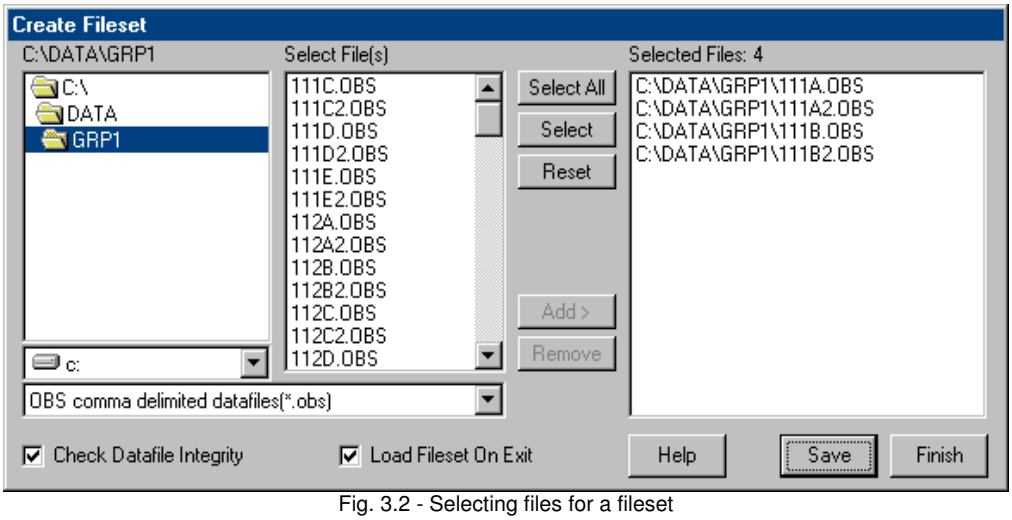


Fig. 3.2 - Selecting files for a fileset

Two or more files must be selected from the list of datafiles. Select files from the centre window and click **Add >** to place them on the selected file list (files on the list can be deselected by highlighting them and clicking **Remove)**.

All highlighted files may be deselected by clicking **Reset**.

Checking the **Checking Datafile Integrity** box checks each of the selected datafiles before they are saved to the fileset. This ensures that the datafile formats are compatible with ObsWin and that they are not corrupted in any way.

Checking the **Load Fileset On Exit** box automatically opens the newly created fileset into the main window for analysis.

The same format dialogue box (figure 3.2) is used for combining/appending files - see section 4.10. The user can browse directories for different file types (which appear in the centre window), and select files according to given criteria.

File Selection Criteria

When selecting datafiles for creating a fileset, combining or appending, or for inter-observer reliability (section 7) the user may enter selection criteria by clicking the **Select** button to select specific datafiles.

Valid Criteria:-

- `*.*` - all files
- `*.obs` - all files with specified file extension
- `filename.*` - all files of filename with any extension
- `*name*` - all files with the search criteria string somewhere within the file name.

There are also **Select All** and **Reset** buttons.

When the required files are highlighted click **Add >** to add them to the list (deselect files with **Remove**).
The new fileset/datafile will require a new name if it is to be saved. Clicking **Save** will prompt for this, and **Finish** will close the dialogue.

Filesets appear as a table listing the content files. One may double click on any datafile line to open an individual dataset - the fileset will be still present in the background and closing the individual datafile will return to it.

The fileset will be saved with the *.FIL extension.

Filesets can also be saved/renamed by choosing

File
→ **Save Fileset As**

3.5 Load/Save Settings

Choosing

File
→ **Load Settings…**
or → **Save Settings…**

Loads or saves an environment settings (*.STF) file. An environment settings file contains information such as the variables used for data collection, graph colours etc., and can be useful if multiple configurations are frequently used.

3.6 Closing Files

Choosing

File
→ **Close Datafile**
or → **Close Fileset**

closes the datafile (or fileset) and prompts to save the data if any edits have taken place.

4. Editing Data

One can view a datafile for editing by choosing

File
→ **Open**
→ **Datafile**

or clicking  on the main toolbar. Each line in the grid represents an observation: its variable name, onset time and offset time respectively.

The grid can be scrolled up and down if it is longer than the window.

In-place editing

- Select the cell where you wish to edit data, either by clicking with the mouse or navigating with the arrow keys.
- Press **RETURN** or double click the cell to activate.
- Type in the new data and either press **RETURN** or click on another part of the window to confirm the change.

At the bottom right of the main screen (see figure 3.1) the following information is displayed: filename, file date and the total file length (in time intervals).

Datafiles can be edited from within ObsWin in a similar way to other spreadsheet packages. A variety of editing features permit almost any manipulation of the dataset. The arrow keys or the mouse can be used to move around the cells. Select cells or rows with the **Shift+Arrow** keys or drag with the mouse. Clicking on an observation number will select the entire row.

4.1 Adding Rows (Observations)

Make sure that the current cell or cells are in the row where you wish to add a line of data (i.e. if the currently selected row is 10 then the new observation will be placed in line 10).

Choosing

Edit
→ **Add Data**

will add a new row at the current position, shifting the rest of the data down one line. The new row will contain the name 'VAR?'.

4.2 Cutting, Copying and Pasting Data

Select a row or rows and choosing

Edit
→ **Copy Data**
or → **Cut Data**

will cut or copy the selected data into memory (the Windows clipboard).

By highlighting the destination row and choosing

Edit
→ **Paste Data**

the data will be inserted at the current position.

Note ObsWin will detect any out of sequence variables created in this way and warn the user when any procedure is attempted.

4.3 Deleting Rows (Observations)

Highlight the row or rows where you wish to delete data - data are removed by row, not by individual cell - and either press **DEL** or choose

Edit
→ **Delete Data**

Note You need not select all the cells to delete, simply selecting the cells in one column and pressing **DEL** will delete all data in the corresponding rows.

4.4 Deleting Variables

Choosing

Edit
→ **Delete Variables**

allows the user to delete all instances of the specific variable(s) selected from the datafile(s).

4.5 Creating New Variables

When data are to be analyzed, it is sometimes useful to group a number of variables into one composite variable, or create new variables from existing ones.

Note Any newly created variables are added to the dataset at the appropriate time intervals, rather than replacing the component variables. They can be analyzed and saved with the datafile.

(a) From Variable Combinations

There are a number of ways to create composite variables:

Choosing

Edit
→ **Create Variables**
→ **Variable Combinations**
→ **New Combination...**

will open the following dialogue box:

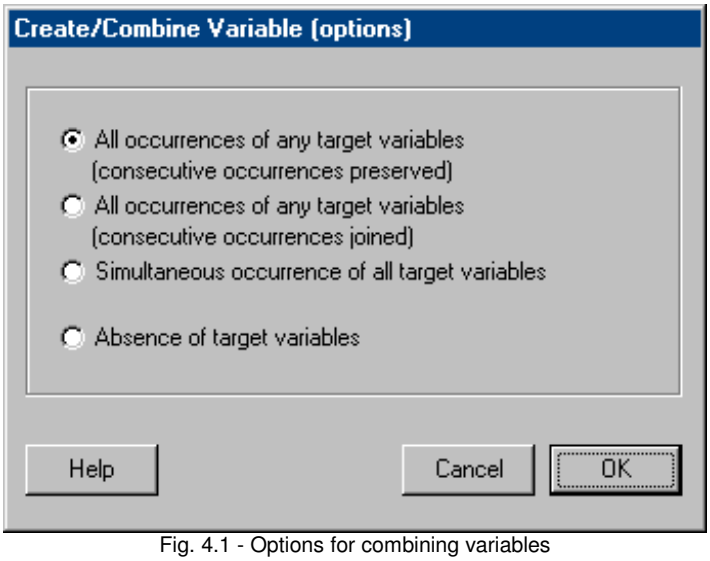


Fig. 4.1 - Options for combining variables

There are four options which generate different information. Make a choice and click **OK.**
A '**Select Variables**' box will appear. Highlight the variables to combine.

To understand how the selections work, think of Boolean logic (*AND*, *OR*, *NOT*). Each choice combines existing variables in one of these ways.

All Occurrences of any Target Variables (Consecutive Occurrences Preserved)

Selecting this option combines any overlapping occurrences of any of the specified target variables as the new combination variable.

All Occurrences of any Target Variables (Consecutive Occurrences Joined)

This option is similar to the one above (Consecutive Occurrences Preserved) except that if successive occurrences of variables are separated by only one interval then they are combined.

Simultaneous Occurrence of All Target Variables

When all of the specified target variables are co-occurring then the new combination variable will have the start time of the interval where the co-occurrence starts, and the finish time of the interval where the co-occurrence ends.

Absence of Target Variables

When none of the specified target variables are occurring the new variable will represent their non-occurrence.

After target variables have been picked and one of the four combination options has been selected, further combinations may also be selected (the different combinations will all be generated simultaneously). One also has the option to save the combinations options as a combination configuration file (*.CCF) to be applied to subsequent datasets.

To apply a combination configuration file to other datafiles/filesets choose

Edit
→ Create Variables
→ Variable Combinations
→ Apply Configuration File...

(b) Create Variable Bursts

Choosing

Edit
→ **Create Variables**
→ **Create Variable Bursts…**

allows one to create a variable burst (i.e. a grouping of successive occurrences of a specified variable).

The user must specify the variable, a new variable name and a time interval for ObsWin to use (maximum number of intervals between successive occurrences).

For example, given the following variable occurrences

	Onset	*Offset*
SHOUT	5	30
SHOUT	32	50
SHOUT	55	130
SHOUT	134	150

if the burst interval selected was 5 then the new variable would become

NEWSHOUT	5	150

(c) Keyed From Variable Onset/Offset

Choosing

Edit
→ **Create Variables**
→ **Keyed From Variable Onset…**
or → **Keyed From Variable Offset…**

allows one to create a variable that is keyed from (contingent on) the onset/offset of another criterion variable.
The user must specify the criterion variable, a new variable name and the size (in time intervals - *n*) of the new variable bout length.
New variables are created either keyed to onset (*n* intervals immediately prior to onset) or keyed to offset (*n* intervals immediately following offset) of the specified criterion variable.

Note Where the interbout between occurrences of the criterion variable is less than the specified bout length then the new variable bout length will be the criterion interbout length.

4.6 Renaming/Labelling Variables

Variables can be labelled after data collection, and existing variable names can be changed.

To rename all occurrences of a variable, choose

Edit
→ **Rename Variables**
→ **Edit Variables…**

Double click on the name in the **New Definition** column and enter the new text. Click on **OK** to update the file.

In the **Rename Variables** dialogue box there are a number of buttons which allow the user to save the new definitions as a Variable Template (*.VTF) file, apply an existing template, and add/delete variable names.

Note Variable template files (*.VTF) contain a set of old variable names and their respective new names, and can be used to automatically rename the same set of variables in the future.

Choosing

Edit
→ **Rename Variables**
→ **Apply Template…**

allows one to rename variables with a template directly.

Variable Lists

To create a variable list simply select

Edit
→ **Create Variable List…**

Choose the number of variables and enter the names manually.

Variable list files (*.VAR) permit the user to load in the same variables (in the same order) when selecting multiple variables. Many analysis options display the **Use List** button when selecting variables, and ObsWin will prompt for a choice of *.VAR files if selected.

Searching for Variables

Choosing

Edit
→ **Find Variable…**
or → **Find Next**

allows the user to locate any variable in the file. Highlight the variable name and click **OK**.

Pressing **F3** (**Find Next**) allows the user to find the next variable searched for in the file.

4.7 Offsetting Data

Choosing the menu option

Edit
→ **Offset time…**

will move every time interval in the file up or down by the specified amount. When more than one person collects data for the same session they are likely to start the timer slightly out of synchrony. This option can be used to synchronise the start times on such recordings for inter-observer reliability. A temporary version of this transformation is available for reliability analysis (see Section 7).

4.8 Splitting Datafiles

Choosing

Edit
→ Split Datafile…

will split the datafile in half (approximately), creating two smaller files.

This may be necessary if the datafile is too large to analyze - this is unlikely, but is a danger if the datafile has more than 10,000 datapoints and/or if it consists of more than 8 hours of data. ObsWin will warn the user if it cannot use all of the data for analysis and will give the option to use the maximum amount of data possible. Splitting will also speed up analyses of large files as it reduces the memory needed by the processor.

The user will be prompted to provide new filenames for the two halves. The smaller files will then be saved and can be analyzed separately.

4.9 Partial Interval Conversion (changing the unit of analysis)

ObsWin's interval recorder is accurate to one second. Partial interval conversion will change the data to reflect a different interval (5, 10, 15 seconds etc.) and the onset and offset intervals will correspond to the new choice (see table 4.1).

Select

Edit
→ Partial Interval Conversion…

and enter the new interval.

Example:-

	Real Time Data 1 Second Interval		*Partial Interval Data 10 Second Interval*	
Variable	*Onset*	*Offset*	*Onset*	*Offset*
A	2	5	1	1
A	15	23	2	3
B	31	52	4	6

Table 4.1 - Partial interval conversion

In this example, the right column shows the recording as if each occurrence of a variable were placed into a 10-second interval rather than one-second. The nature of the conversion will not be marked on the converted datafile (table 4.1 is for illustration only) but the new total number of intervals will appear at the bottom right.

Note Partial Interval Conversion can only 'convert up'. ObsWin does not keep a history of transformations, so data cannot be transformed from 10-second back into 1-second intervals. Look at the column on the right: entering a new interval of 1 would leave it unchanged, entering a new interval of 10 would transform again - to the equivalent of using an interval of 100 on the original datafile.

4.10 Combining and Appending Datafiles

There are two ways to add separate datafiles together to form one datafile. The method used depends on the purpose of joining the files.
Joining files in this way is not the same as creating filesets (see section 3.4).

(a) Combining

Choosing

Edit
→ **Combine Datafiles…**

allows the onset and offset times of the separate files to be interleaved (preserved and sorted) so that the resulting datafile retains the temporal sequence of variables.
However, it is important to remember that onsets and offsets are simply occurrences within a time interval and the precise sequence of variables within a single time interval cannot be guaranteed.

Combining datafiles is useful if observing/recording data from videotape where a number of different variables can be recorded on different runs and the files later combined.

Note This is not recommended since it is unlikely that the user will be able to start the second pass at precisely the same time as the first.

(b) Appending

Choosing

Edit
→ **Append Datafiles…**

allows datafiles to be 'stacked' in the order in which they are selected. This is useful when joining sessions that were continuous, but broken into smaller observation sessions (for whatever reason).

Note Analyzing an appended datafile will not necessarily produce the same results as analyzing each of the files in a fileset as the appended file will not contain the original end of file markers, and is not recommended.

---

5. Graphing data

5.1 Occurrence Graph

Occurrence graphs show the presence and absence of variables over the time of the recording. Multiple variables can be displayed on one graph, stacked one on top of another.

Choose

Graph
→ **Occurrence Display…**

or press **F5**, or click the  button on main toolbar.

Select the variables to be displayed on the graph from the variable list display, and click **OK**. One must then select from a number of **Occurrence Graph Options** (see figure 5.1 below).

Occurrence graph options

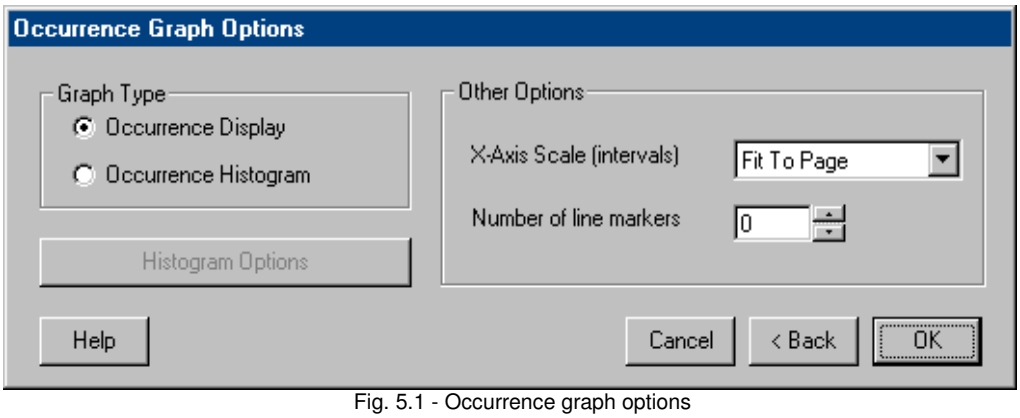


Fig. 5.1 - Occurrence graph options

Graph Type

Selecting an **Occurrence Display** (the default display - figure 5.2) will show an interval by interval occurrence of each of the variables selected.

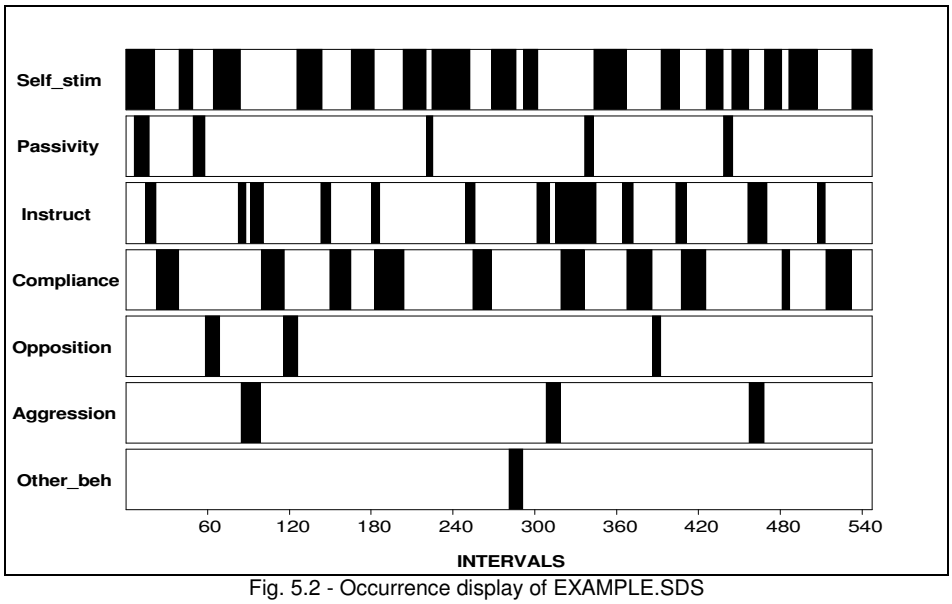


Fig. 5.2 - Occurrence display of EXAMPLE.SDS

Selecting an **Occurrence Histogram** (see figure 5.3) will show the proportion of one second intervals in each 10 second time period for the occurrence of each variable (the default setting). These settings (the number of intervals and the time period) can be changed by clicking the **Histogram Options** button.

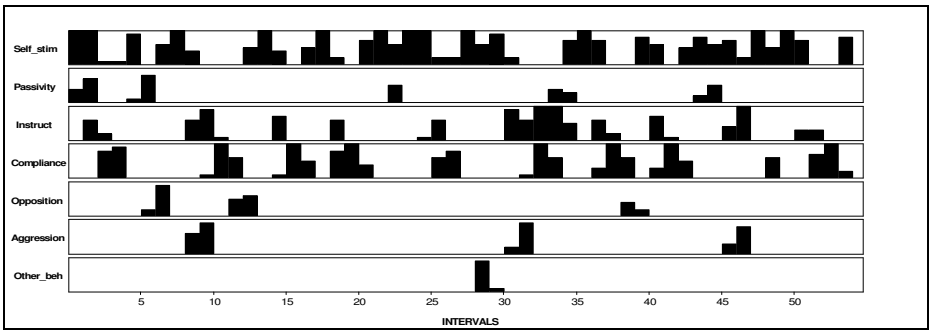


Fig. 5.3 - Occurrence histogram for EXAMPLE.SDS

Note An Occurrence Histogram is similar to an Occurrence Display except that the use of a time period (minimum 10 seconds) effectively displays the data as if recorded using partial interval recording.

Other Options

X-Axis Scale

The default setting for the X-axis scale is **Fit To Page**, which shows all the data on screen at any one time. Alternatively, the user may select a range of intervals from 300 to 6000 (figure 5.2 and 5.3 use different X-axis scales).

Number of line markers

Line markers are vertical dotted lines that can be displayed on the graph at the time indicated with associated text. Any notes or conditions entered during the data collection stage (see section 2.3) will be treated as markers and will be shown prior to the graph being displayed, for annotation and confirmation.

The Graph Menu

Choosing

Graph
→ **Print Graph**

allows output to a printer. For long Occurrence Graphs the graph will be automatically scaled to fit onto one page.

Choosing

Graph
→ **Save Graph…**

allows one to save the graph as either a Windows Metafile (*.WMF) or Windows Bitmap (*.BMP) file.

Choosing

Graph
→ **Copy to Clipboard**

copies the graph to the Windows clipboard which allows it to be pasted into other applications.

Options menu

When an occurrence graph is displayed, the options menu will contain two special choices not present on other graphs:

Choosing

Options
→ **Binary Interval Output…**

allows the occurrence graph data to be saved in binary form as an ASCII file and displayed in the Obs Editor. A '1' indicates occurrence in an interval and '0' (zero) a non-occurrence. This allows the graph data to be used to reconstruct an Occurrence Graph in another graphics application.

Choosing

Options
→ **Occurrence Graph Options…**

allows the user to return to the options menu to change the various options (see figure 5.1).

Options
→ **Graph Settings**

allows the user to change various graphics settings (also available from the main **Options** menu). **Colours** allows the user to customize the colour scheme for any aspect of the graphs, **Axis** switches the inner frame on and off, and allows the zero Y-axis to be raised, **Markers** allows various aspects of data point markers to be changed and **Text** allows the selection of alternative fonts and font size.

5.2 Cumulative Onset Graph

Variables may be displayed as cumulative records which allows one to inspect the scheduling of the variable and also whether there is evidence of a contingent influence of a second variable on the first variable.
Each onset of the first variable (the criterion) is plotted cumulatively (as a line) across time, and the occurrence of a second variable (the target) is marked as a small line.

Choosing

Graph
→ **Cumulative Onset…**

will display a variable list box from which the user must select a criterion and one or more target variables (see figure 5.4 below) by selecting each variable and clicking **Add >** .

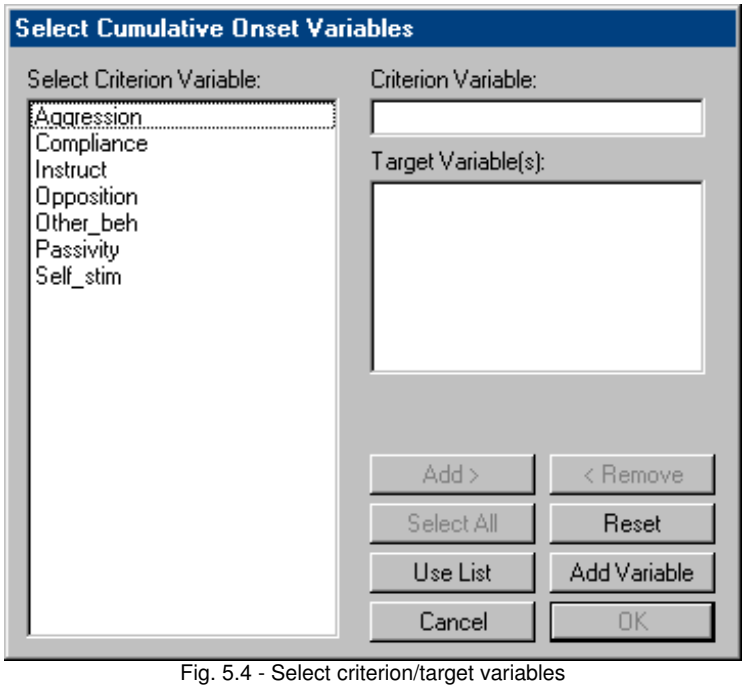


Fig. 5.4 - Select criterion/target variables

Figure 5.5 below shows a cumulative onset graph with the criterion “Instruct” and target “Compliance” variables from the EXAMPLE.SDS datafile.

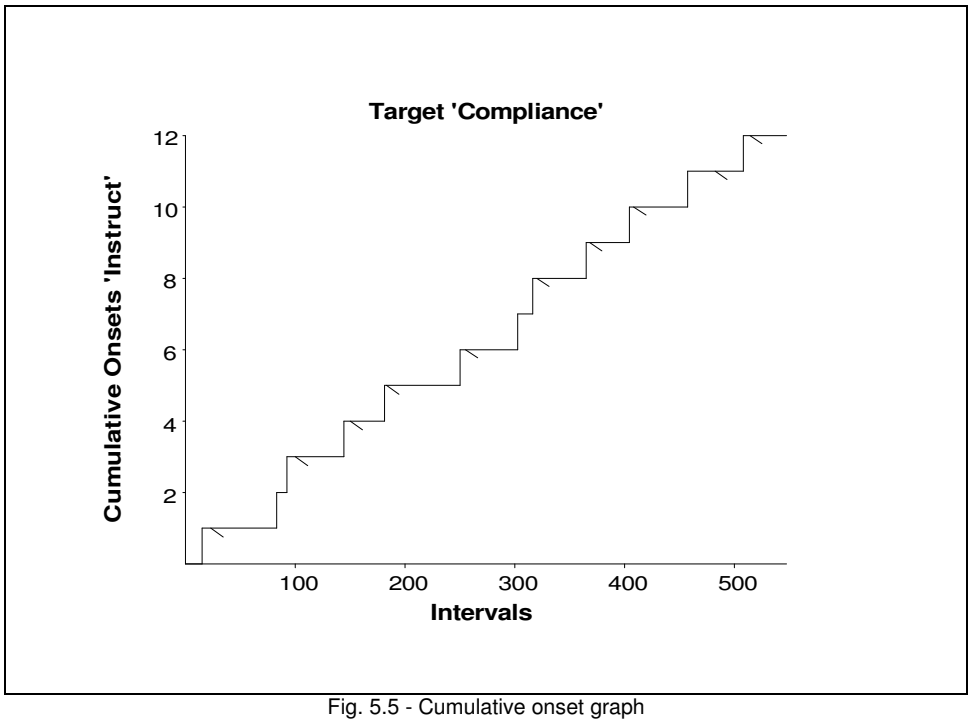


Fig. 5.5 - Cumulative onset graph

If more than one graph is generated, then the user has the option to choose how many of the graphs to display on screen (or print).

Choose

Graph
→ **Select Graph(s)…**

Up to 9 graphs can be displayed on screen or printed at any one time.

Double clicking on any of the graphs displayed will reveal a pop-up menu that allows the user to inspect the raw graph data, print or redisplay the selected graph.

5.3 Frequency Per Interval Graph

A Frequency Per Interval graph plots the frequency with which the selected variable occurs within a specified interval.

Choose

Graph
→ **Frequency Per Interval…**

and select the variables to graph from the variable list shown. The user must then select a time interval.
The data are then broken up into intervals starting from the first interval to the end of the datafile (e.g. if an interval of 10 is used then the intervals are 1 - 10, 11 - 20, 21 - 30 etc.).

Note Where there is an uneven interval left at the end of the datafile, the frequency scores will be prorated.

Figure 5.6 shows Frequency Per Interval graphs (60-second intervals) from the EXAMPLE.SDS datafile.

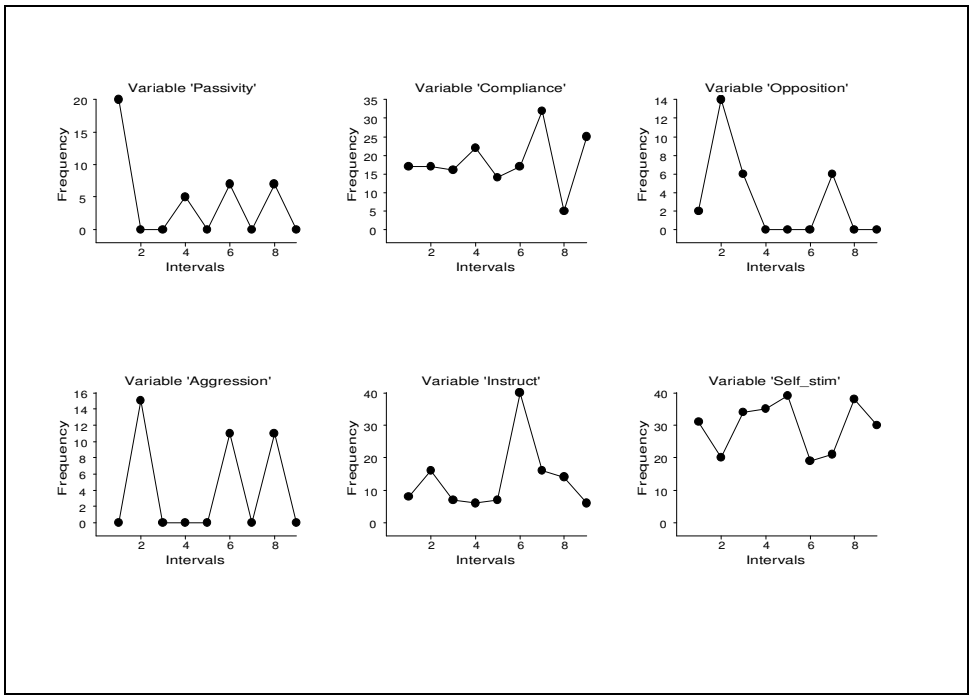


Fig. 5.6 - Frequency per 60 second interval

There are a number of options on the graph menu when displaying a Frequency Per Interval graph.

Choosing

Options
→ **Maximize Y-axis**

Sets the last point on the axis to the maximum Y value.

Choosing

Options
→ **Standardise Y-axis**

sets all the Y-axes to the same scale if one is displaying more than one graph.
Choosing

Options
→ **Show Percentage of Interval**

changes the Y-axis units from a frequency count to a percentage of interval.

5.4 Summary Statistics Bar Charts

Summary statistics bar charts show percentage of interval occurrence for variables.

Choosing

Graph
→ **Summary Statistics Bar Charts**
→ **Each File…**

from the main **Graph** menu, or clicking  on the main toolbar menu, graphs each datafile separately (see figure 5.7).

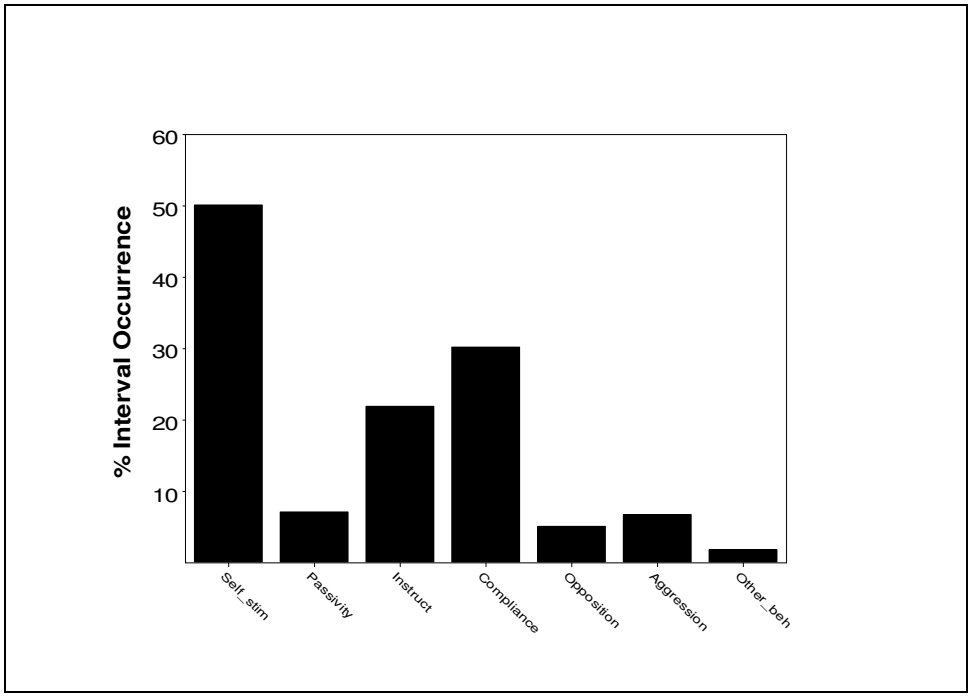


Fig. 5.7 - Summary statistics bar chart (EXAMPLE.SDS)

When graphing multiple files (i.e. from a fileset) then the last three menu options become available.

Choosing

Graph
→ **Summary Statistics Bar Charts**
→ **Totals Across Files…**

sums data across each file and presents total data in one graph.

Choosing

Graph
→ **Summary Statistics Bar Charts**
→ **Mean Across Files…**

calculates the mean data for each variable across files and presents it in one graph.

One may optionally display error bars to show variance for a Mean Across Files graph by choosing

Options
→ **Error bars**

from the Graph display **Options** menu.

Choosing

Graph
→ **Summary Statistics Bar Charts**
→ **SD Across Files…**

calculates mean data for each variable, and displays each variable on a separate graph showing the standard deviations from the mean. Datafiles are marked on the X - axis.

Note Clicking on the  button on the main toolbar or pressing the **F6** key will repeat the last chosen graph option using the same chosen variables.

5.5 View/Print Graphs

Graphs that have been saved as Windows Metafiles (*.WMF) or Windows Bitmaps (*.BMP) may be viewed and printed by choosing

Graph
→ **View/Print Graph…**

One may add titles/headers to graphs to print by choosing

Graph
→ **Graph Header…**

from the View/Print Graph menu.

6. Analyzing data

6.1 Summary Statistics

Choosing

Analyze
→ Variable Summary Statistics…

or pressing **F7**, or clicking the  button on the main toolbar provides descriptive statistics for any number of variables.

In the **Select Variables for Summary Statistics** box, highlight the required variables and click **OK** to continue.

If using a fileset ObsWin will offer 3 further **Summary Statistics Options**:

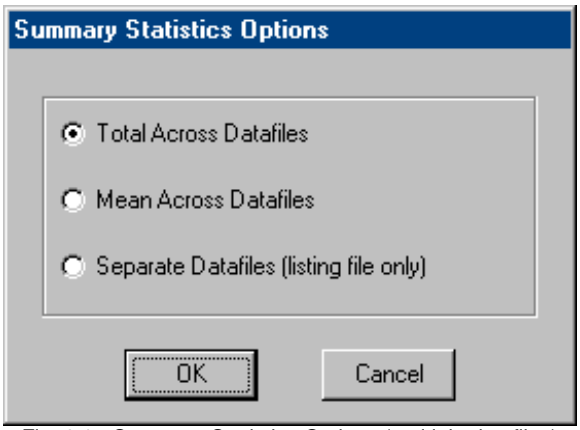


Fig. 6.1 - Summary Statistics Options (multiple datafiles)

Selecting, **Total Across Datafiles** sums data across each file, selecting **Mean Across Datafiles** presents mean data across the files and selecting **Separate datafiles** sends the output to the listing file only.

Summary statistics appear in a separate data display grid (see figure 6.2).

Summary Statistics - Example.sds (Total intervals = 547)

Variable	Type	Intervals	% Intervals	Mean Duration	Median Duration	Onsets	Offsets
Self_stim	Duration	274	50.09	17.13	17	15	15
Passivity	Duration	39	7.13	7.80	9	5	5
Instruct	Duration	120	21.94	10.00	8	12	12
Compliance	Duration	165	30.16	16.50	17	10	10
Opposition	Duration	28	5.12	9.33	11	3	3
Aggression	Duration	37	6.76	12.33	11	3	3
Other_beh	Duration	10	1.83	10.00	10	1	1

Fig. 6.2 - Summary statistics output (EXAMPLE.SDS)

Data are presented in columns as follows:

Type: indicates whether a variable is a duration or event variable, or “unknown” (i.e. does not occur in the datafile).
Intervals: the number of intervals in which the variable occurs.
Percentage Of Intervals: the previous column as a proportion of the total number of intervals.

Mean Duration: the mean duration for each bout of a variable.
Median Duration: the median duration for each bout of a variable.
Onsets/Offset: the number of onsets and offsets of a variable: a discrepancy indicates a variable in progress at the start or finish of the recording (see **Analysis Options** section 6.9).

A number of options are available from the Summary Statistics **Data** menu.

Choosing

Data
→ **Graph Data**

displays the percentage of intervals in which the variable was occurring.

Note The graph displayed will be the same as the graph displayed when choosing **Summary Statistics Bar Charts** from the main graph menu.

Choosing

Data
→ **Save Spreadsheet Data…**

saves the summary statistics data in tab delimited format, which can be imported into other spreadsheet applications such as Microsoft Excel.

Choosing

Data
→ **Frequency Analysis…**

provides frequency scores per *n* intervals (user defined) and appends them to the summary statistics table. The default number of intervals is 60 (i.e. the frequency analysis will give a frequency per minute).

6.2 Bout/Interbout Distributions

Choosing

Analyze
→ **Bout/Interbout Distributions…**

generates a graphical display of both bout and interbout frequency distributions, one graph for each (see figure 6.3) - the raw data (frequencies of bouts/interbouts) are saved to the listing file.

The default graph type produced is a line graph.
Choosing

Options
→ **Bar Chart**

changes the graph type to a bar chart.

Choosing

Options
→ **Show as Percentages**

changes the Y-axis display units from frequency to percentage of the total frequency.

Choosing

Options
→ Limit X-Axis Scale…

allows the user to limit the X-axis on any of the graphs displayed.

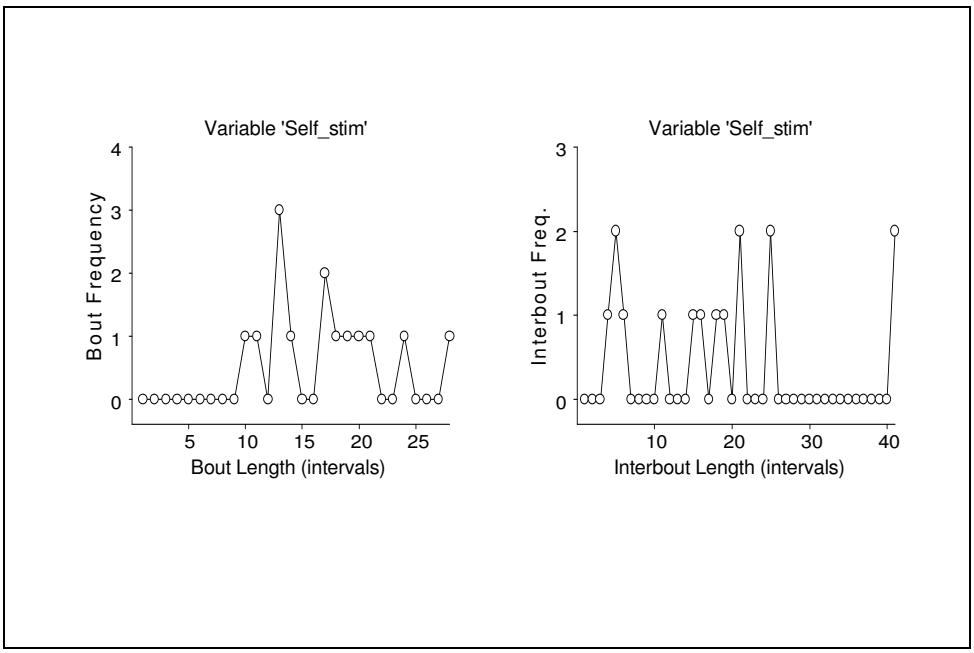


Fig. 6.3 - Bout/interbout distributions - EXAMPLE.SDS

6.3 Markov Chain Analysis

A Markov Chain analysis determines whether a group of variables occur together in a sequence (of event units) more often than would be expected by chance.
In addition it determines whether the dataset consists of significant chains of 2, 3 or 4 variables.

Choose

Analyze
→ Markov Chain Analysis

and select two or more variables and click **OK**.

An options dialogue box is then presented as follows:

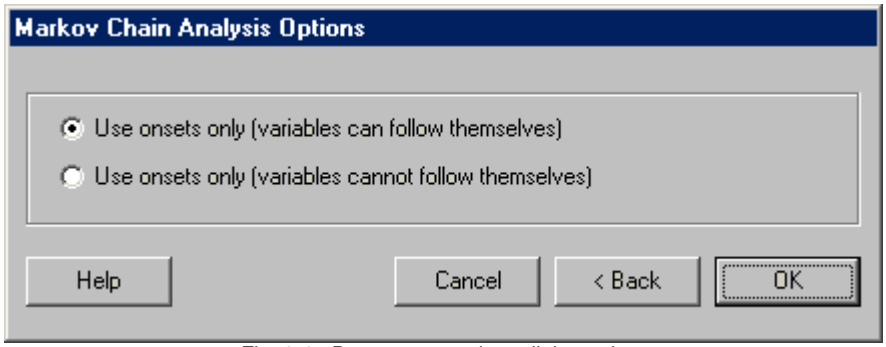


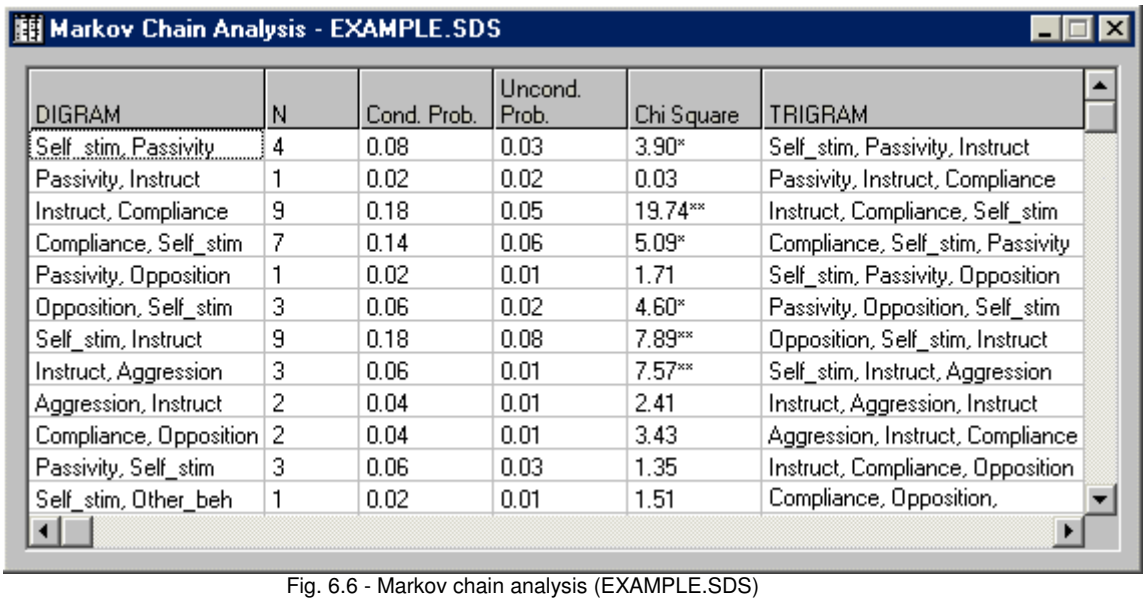
Fig. 6.4 - Datastream options dialogue box

Selecting **Create Mutually Exclusive/Exhaustive Variables** will re-order the datastream to take account of variable durations and gaps in the dataset and as well as co-occurring variables. Gaps are treated as variables in their own right (denoted by the variable name *< GAP >*) and co-occurring variables will become separate variables in the analysis and are denoted by the plus sign (e.g. *VAR1+VAR2*).

Selecting **Use Onsets Only (Variables Cannot Follow Themselves)** will treat each new variable onset as the next variable in sequence, although if the same variable follows itself it will be discounted.

Use Onsets Only (Variables Can Follow Themselves) is the same as the second option except that a variable can follow itself.

The output data for a Markov Chain Analysis are displayed in a grid (see figure 6.6) consisting of conditional and unconditional probabilities for each different sequence of *n* variables, although only for sequences that occur more than once, and a Chi Square value to indicate whether the difference between the unconditional and conditional probability is significant. The grid is scrollable, showing Digrams on the left (all possible chains of 2 variables) followed by the Trigrams (chains of 3) and Tetragrams (chains of 4) on the right. Chains longer than 4 are not displayed.



DIGRAM	N	Cond. Prob.	Uncond. Prob.	Chi Square	TRIGRAM
Self_stim, Passivity	4	0.08	0.03	3.90*	Self_stim, Passivity, Instruct
Passivity, Instruct	1	0.02	0.02	0.03	Passivity, Instruct, Compliance
Instruct, Compliance	9	0.18	0.05	19.74**	Instruct, Compliance, Self_stim
Compliance, Self_stim	7	0.14	0.06	5.09*	Compliance, Self_stim, Passivity
Passivity, Opposition	1	0.02	0.01	1.71	Self_stim, Passivity, Opposition
Opposition, Self_stim	3	0.06	0.02	4.60*	Passivity, Opposition, Self_stim
Self_stim, Instruct	9	0.18	0.08	7.89**	Opposition, Self_stim, Instruct
Instruct, Aggression	3	0.06	0.01	7.57**	Self_stim, Instruct, Aggression
Aggression, Instruct	2	0.04	0.01	2.41	Instruct, Aggression, Instruct
Compliance, Opposition	2	0.04	0.01	3.43	Aggression, Instruct, Compliance
Passivity, Self_stim	3	0.06	0.03	1.35	Instruct, Compliance, Opposition
Self_stim, Other_beh	1	0.02	0.01	1.51	Compliance, Opposition,

Fig. 6.6 - Markov chain analysis (EXAMPLE.SDS)

6.4 State Transitions

State transitions are a form of lag sequential analysis. Unlike a Markov Chain analysis, a State Transition analysis determines the likelihood of one variable occurring given the prior occurrence of another. Data can be analyzed in two units: event units or time intervals. For an analysis conducted in event units, conditional probabilities reflect the extent to which one variable precedes or follows another, irrespective of the number of time intervals occurring between the two variables. For an analysis conducted in time interval units, the conditional probabilities reflect the extent to which one variable precedes or follows another variable in time.

The user is required to select a minimum of two variables. If more than two are chosen, all permutations for pairs of variables will be analyzed.

To conduct the analysis in event units, choose

Analyze
→ State Transitions
→ Event…

and select variables for the analysis.

If only some of the variables are picked from the list then one is presented with a list of **Variables Of Interest** to select (see figure 6.7 below).

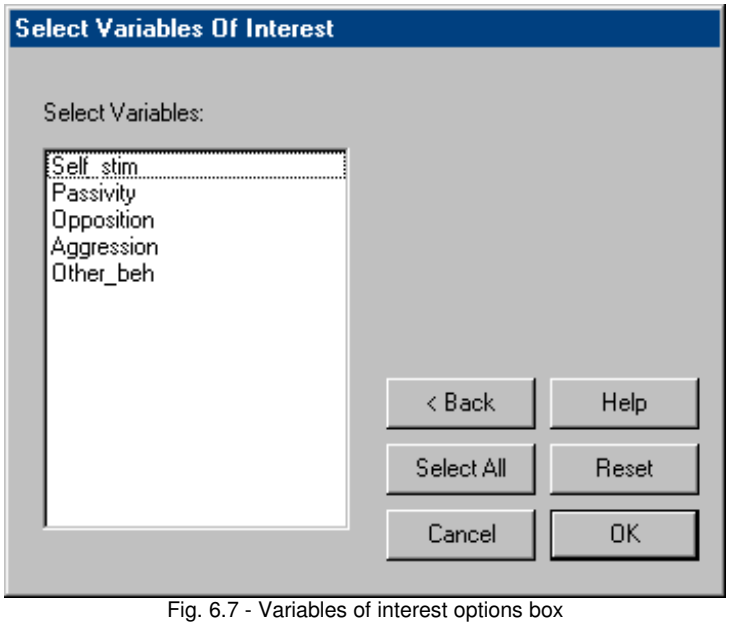


Fig. 6.7 - Variables of interest options box

Variables of Interest are used to tell ObsWin that something is happening in between the variables being analyzed, or else it assumes that nothing is happening.

Figure 6.8 shows an event-based state transition diagram for the variables “Instruct” and “Compliance”, with conditional probabilities for each possible state transition.

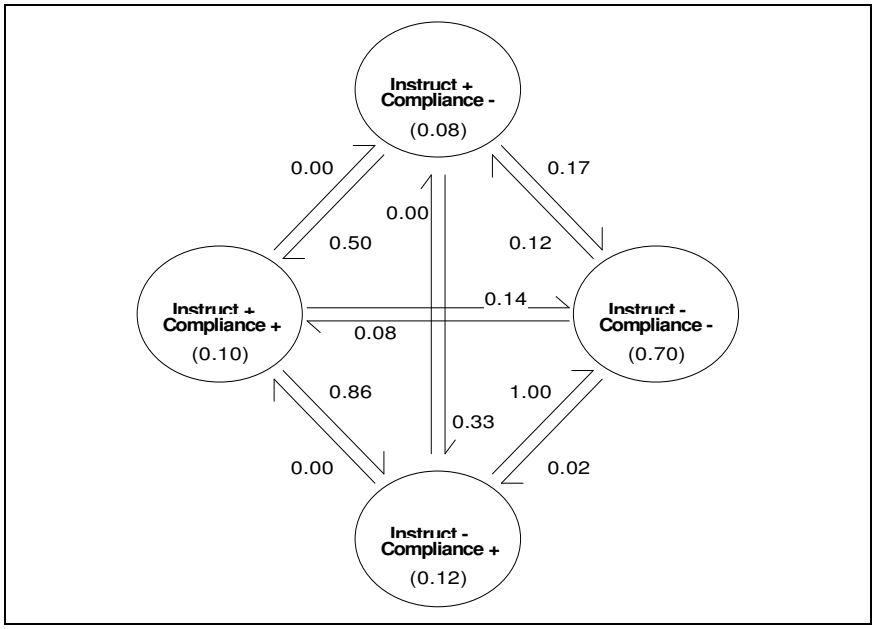


Fig. 6.8 - Event-based state transition diagram

To conduct the analysis in time units, choose

Analyze
→ State Transitions
→ Time…

Select two or more variables and the following options will appear:

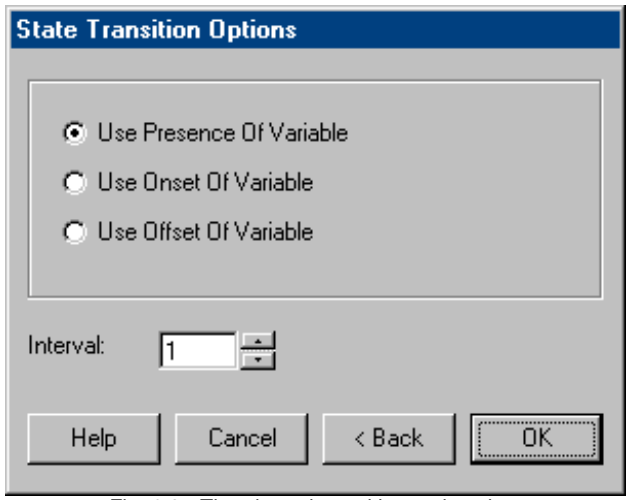


Fig. 6.9 - Time-based transition options box

Select **Use Presence Of Variable** to count every interval in which the variable is occurring. Select **Use Onset Of Variable** to count only the time intervals at which the onset of the variable occurs. Select **Use Offset Of Variable** to count only the time intervals at which the offset of the variable occurs.

The default interval for the analysis is 1. If the interval chosen is greater than 1 then the next n intervals are used to determine the conditional probabilities for each change in state.

For an event-based state transition analysis there are 12 possible transitions and for a time-based one there are 16 (states may follow themselves) - see figure 6.10.

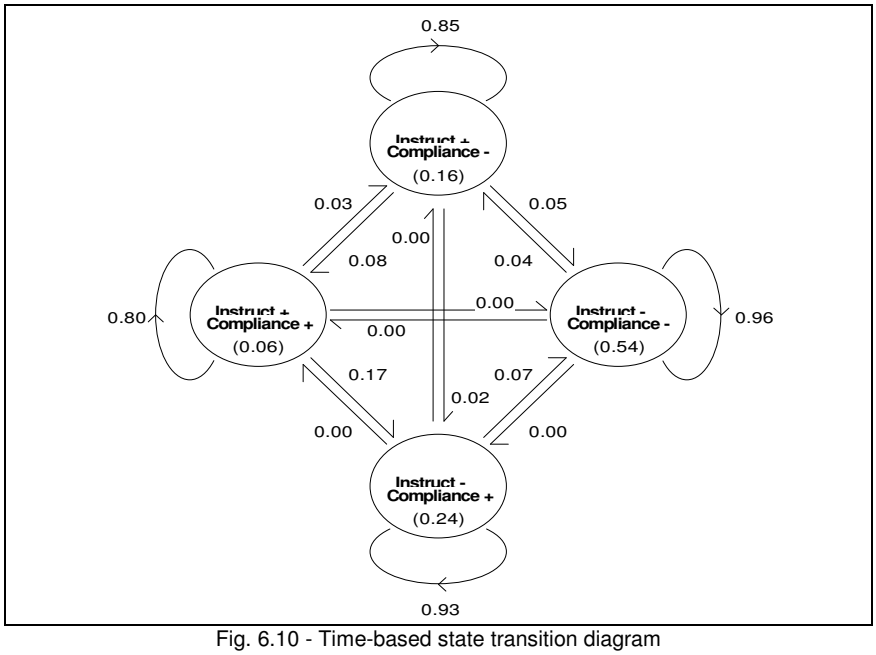


Fig. 6.10 - Time-based state transition diagram

6.5 Autocorrelation

Autocorrelation (sometimes referred to as autolagging) lags a variable against itself using successive events or intervals to determine the conditional probability of its presence at successive lags.

For an event-based analysis choose

Analyze
→ **Autocorrelate Variable**
→ **Event…**

and select variables for autocorrelation (and **Variables of Interest** if applicable **-** see section 6.4).

The same datastream dialogue box appears as that used in the Markov Chain Analysis:

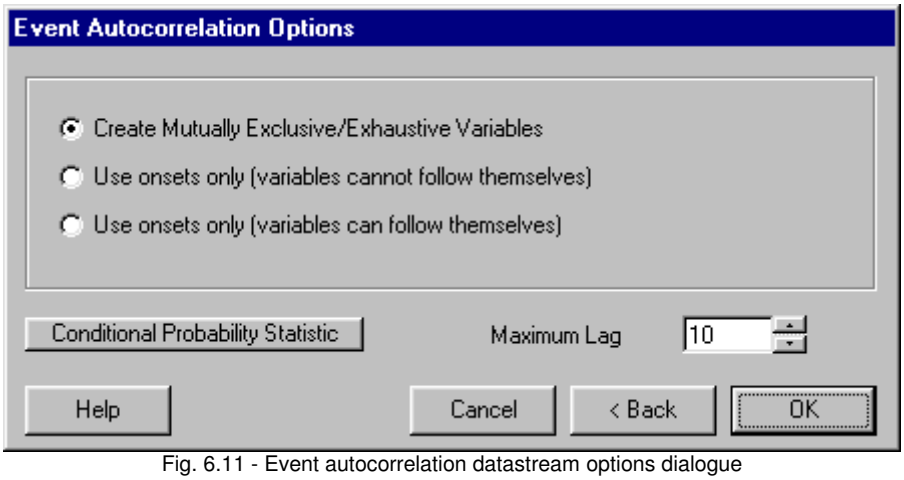


Fig. 6.11 - Event autocorrelation datastream options dialogue

There is one extra option: **Maximum Lag** specifies the upper limit on the number of lags to be considered (default is 10).

Clicking on **Conditional Probability Statistic** will open this window:

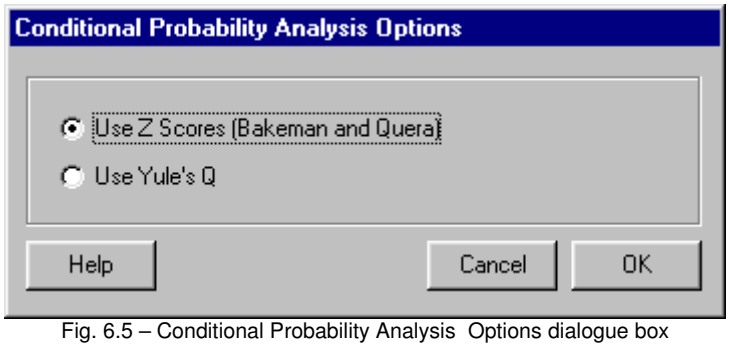


Fig. 6.5 – Conditional Probability Analysis Options dialogue box

Z Scores (Bakeman and Quera)

Z is calculated from the normal distribution approximation to the exact binomial probability, and requires the assumption that the criterion behaviour probabilities are asymptotic population values not subject to sampling variation - this assumption can only really be justified with very large datasets. See references in manual for details on computation (Bakeman & Quera, 1995; Bakeman & Gottman, 1997).

Z Scores of <= -1.96 or >= 1.96 are significant at 0.05.
Z Scores of <= -2.56 or >= 2.56 are significant at 0.01.

Yule's Q

Q is a function of the cross-ratio for a 2 x 2 contingency table and has a range of -1 to +1. Q is computed as (AD-BC)/(AD+BC) where A represents both target and criterion variables occurring, B and C represent either the presence of the target or criterion variable, and D represents non-occurrence of both target and criterion variables.

The output is two graphs, one for probability and one for Z Scores or Q Scores.

Figure 6.12 shows an event-based autocorrelation for the variable “Self_Stim”.

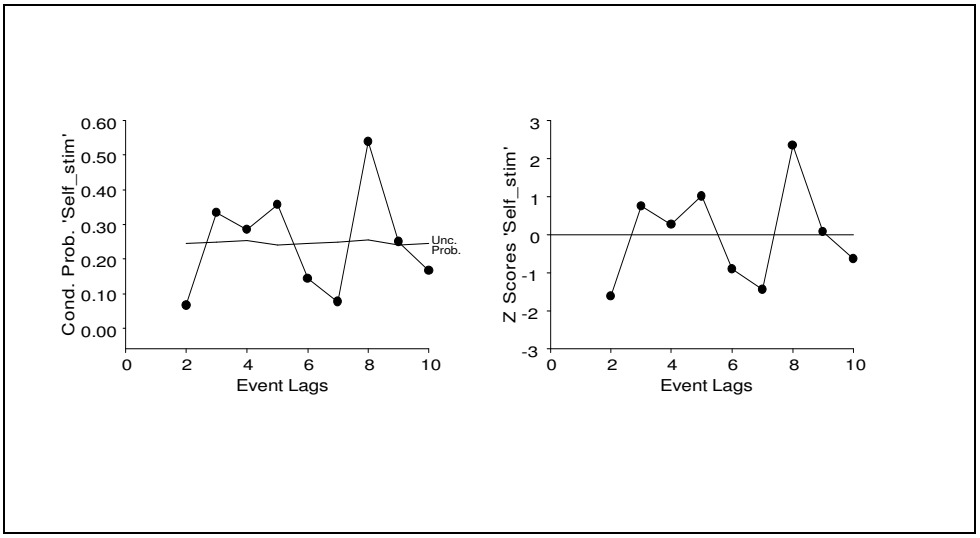


Fig. 6.12 - Event-based autocorrelation

For a time-based analysis choose

Analyze
→ **Autocorrelate Variable**
→ **Time…**

and select variables for the analysis. The following dialogue option box is presented:

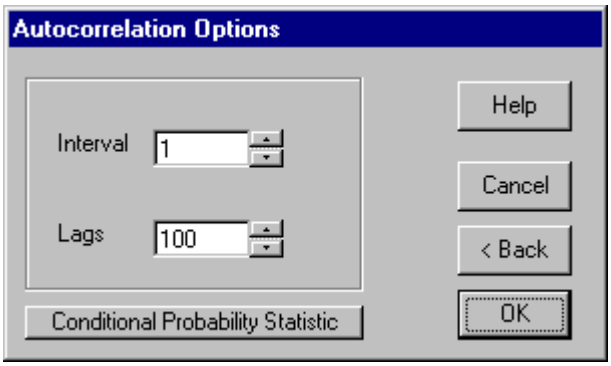


Fig. 6.13 - Autocorrelation options dialogue

The **Interval** option determines the size of the lag to be used (the default is 1 interval) and the **Lags** option determines the number of lags to be plotted on the graph (the default is 100).

Figure 6.14 below shows a time-based autocorrelation graph for the variable “Self_Stim”.

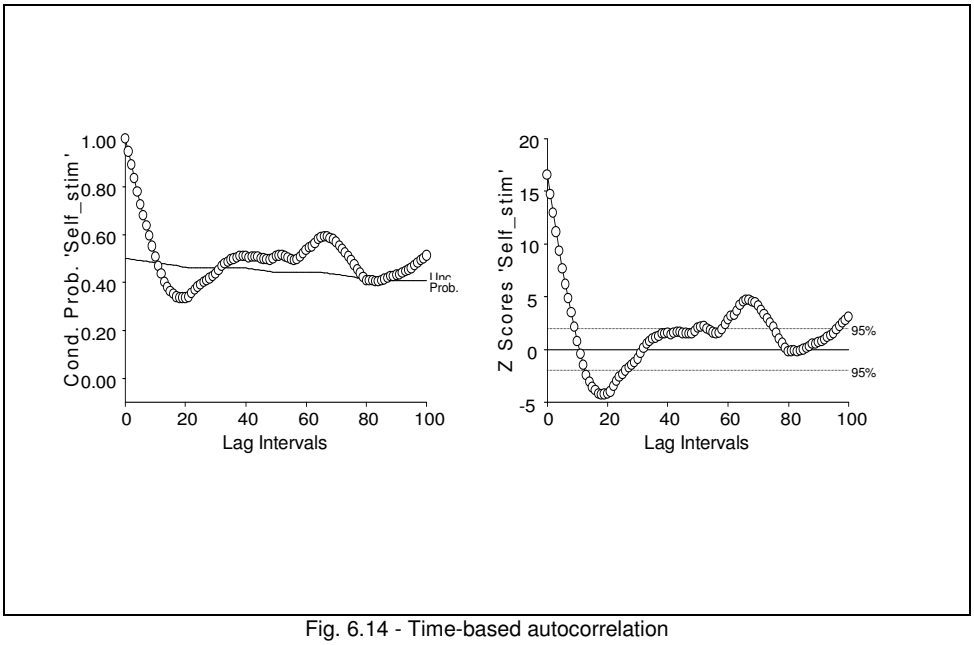


Fig. 6.14 - Time-based autocorrelation

Choosing

Options
→ **Bout Criterion**

from the Graph menu determines the point at which the variable is no longer likely to occur (the point at which it is no longer autocorrelated).

Choosing

Options
→ **Confidence Limits…**

from the Graph menu will mark the 95% and 99% significance intervals on the Z Score graph (see figure 6.14).

6.6 Lag Analyses

Overview

Unlike State Transition analysis, lag analysis can consider more than one lag. Unlike Autocorrelation, it can consider more than one variable.
Lag analyses can be event or time-based.
By convention, if a target variable occurs concurrently with a criterion variable then this is denoted Lag 0 (only possible with time units). If a target variable occurs in the next event or time interval following the criterion, then this is denoted Lag 1 (and if it occurs in the next but one interval then this is denoted Lag 2 etc.). Any events or time intervals preceding Lag 0 are said to occur at Lag -1, Lag -2 etc.

(a) Event-Based Lag Analysis

For an event lag analysis choose

Analyze
→ **Lag Analysis**
→ **Event…**

The user must then select at least 2 variables, one as the criterion (or given variable) and one or more target variables (see figure 5.4).

If not all of the available variables are selected then the user may **Select Variables of Interest** (see figure 6.7).

The following lag option dialogue box will appear:

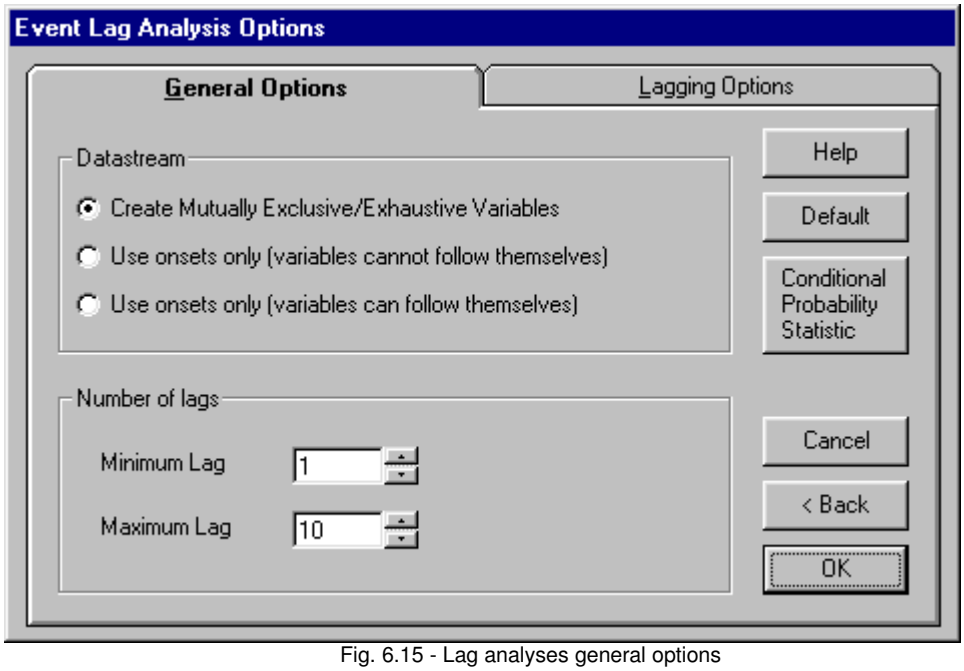


Fig. 6.15 - Lag analyses general options

General Options

The first three **Datastream** options are described in section 6.3.
The **Number of lags** option allows the user to specify the minimum and maximum lags.
The **Minimum Lag s**pecifies the starting point for the lag analysis (e.g. if this is set at -10 then the conditional probability of a target variable occurring 10 event units prior to the occurrence of the criterion variable is calculated - the default minimum lag is 1). The **Maximum Lag** specifies the finishing point for the analysis (e.g. if this is set at +10 then the conditional probability of a target variable occurring 10 event units following the occurrence of the criterion variable is calculated – the maximum lag must be greater than or equal to the minimum lag - the default is 10).

Note The lag range may not exceed 300.

Lagging Options

Clicking on the **Lagging Options** section allows the user to change the default **Restrictions** and **Criterion/Target** settings (see figure 6.16).

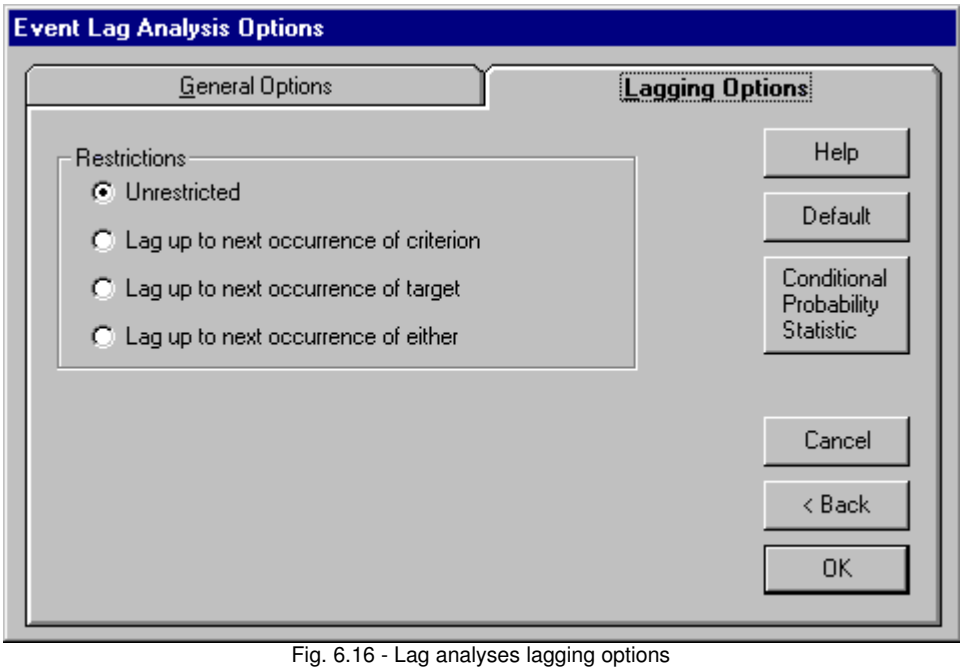


Fig. 6.16 - Lag analyses lagging options

The user may change the default option (**Unrestricted**) by selecting **Lag up to the next occurrence of criterion** (i.e. when the next occurrence of the criterion variable is found, no further lagging occurs), or **Lag up to the next occurrence of target** or **Lag up to next occurrence of either**.

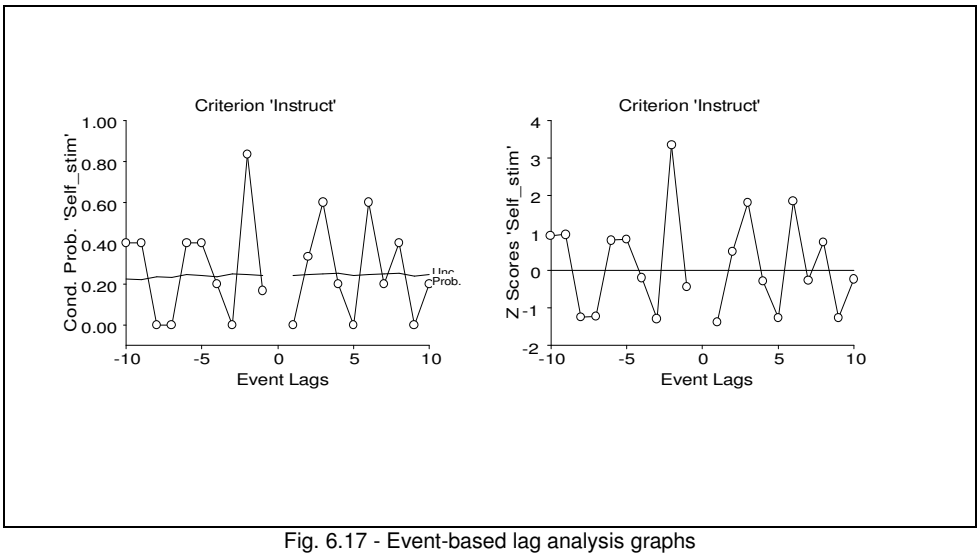


Fig. 6.17 - Event-based lag analysis graphs

Note There will be no lag zero data for an event-based lag analysis as events cannot co-occur.

Choosing

Options
→ **Confidence Limits…**

will mark the 95% and 99% significance intervals on the graph.

(b) Time-Based Lag Analysis

For a time-based lag analysis choose

Analyze
→ **Lag Analysis**
→ **Time…**

and select the criterion and target variables.

The following lag options dialogue box appears:

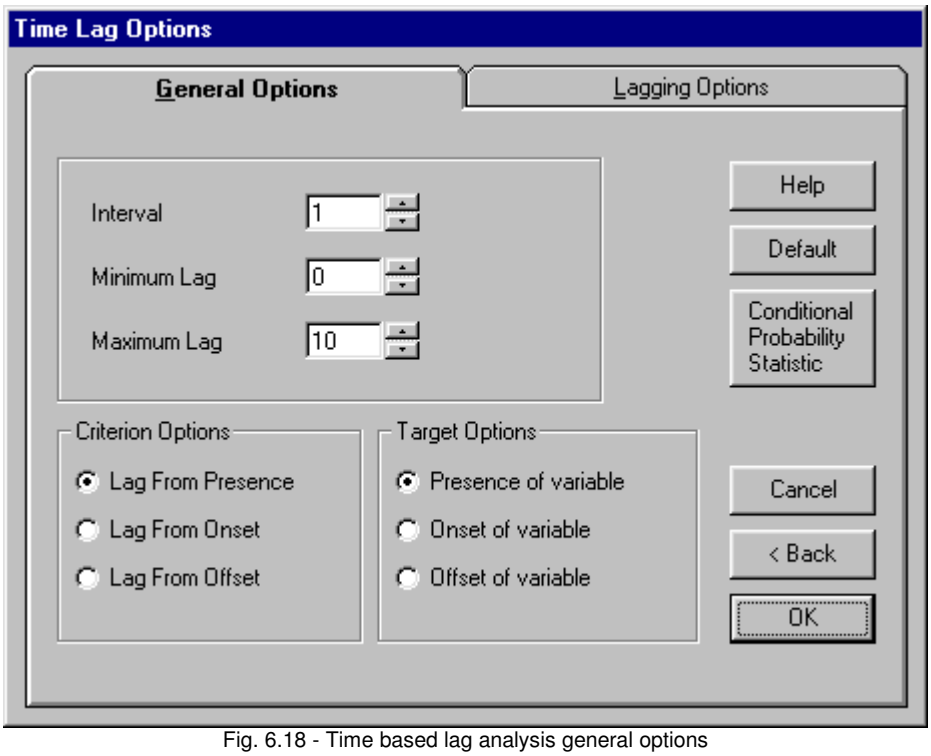


Fig. 6.18 - Time based lag analysis general options

General Options

The **Interval** determines the size of the lag unit (measured in time intervals) - the default is 1.
The Minimum Lag and Maximum Lag determine the starting point and finishing point of the lagging in time units (see Event-Based Lag Analysis above).

The user may select from a number of **Criterion Options**.

The default option, **Lag From Presence**, allows every interval in which the criterion variable is occurring to be used to initiate lagging. Selecting **Lag From Onset** allows only the onset of the criterion variable to be used to initiate lagging. Selecting **Lag From Offset** allows only the offset of the criterion to be used to initiate lagging.

The user may also select from a number of **Target Options**.
The default option, **Presence Of Variable**, counts every interval in which the target variable is occurring. Selecting **Onset Of Variable** counts only onsets of the target variable and selecting **Offset Of Variable** counts only the offsets.

Lagging Options

Clicking on the **Lagging Options** section allows the user to change the default **Restrictions** settings (see Event-Based Lag Analysis - figure 6.16).

The default options will produce a graphical display of the data (see figure 6.19 below).

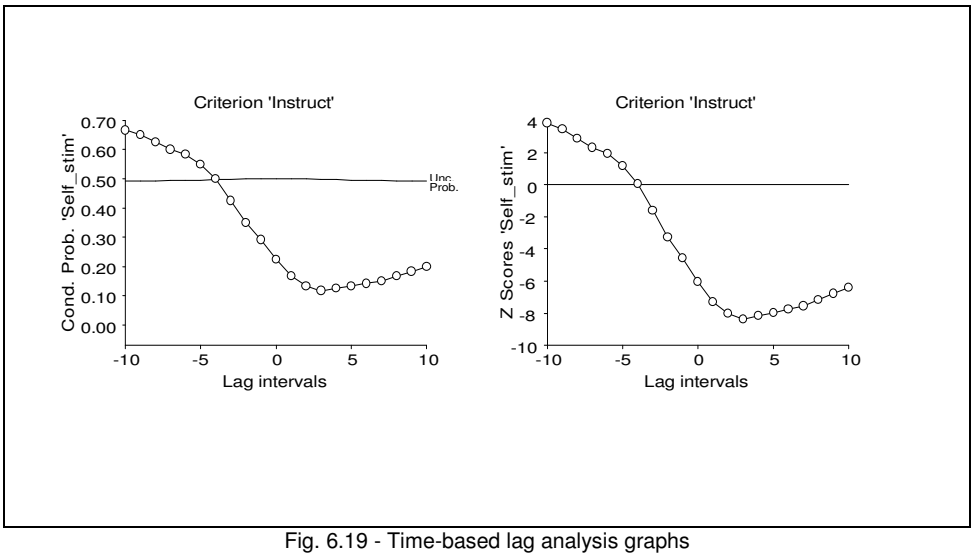


Fig. 6.19 - Time-based lag analysis graphs

6.7 Burst Analysis

Burst analysis extends the principle of lag analysis to allow the temporal association between a criterion and target variable to be examined. Specifically, three time periods are defined - *prior* to the occurrence of a criterion variable, *during* it's occurrence and *following* it's occurrence. The occurrence of a target variable within these three periods is then determined, either in absolute time units (intervals) or relative time units (percentiles).

(a) Burst Analysis (Using Intervals)

Conditional probabilities are calculated (for each interval prior to, during and following each burst) by dividing the number of times the target behaviour has occurred at a specific interval by the number of bursts where the specific interval is still within range (see Emerson *et al*, 1996). Since the bursts are unlikely to be the same length, the denominator will not simply be the number of bursts but will diminish as the burst lengths/interbouts get longer. Because of this ObsWin scans both forwards (i.e. from the beginning to the end of a burst) and backwards (from the end to the beginning) and displays both results to give a fairer representation of the transitions over these periods.

For an interval analysis choose

Analyze
→ Burst Analysis
→ Using Intervals…

and select a criterion and one or more target variables.

The following options dialogue box will appear.

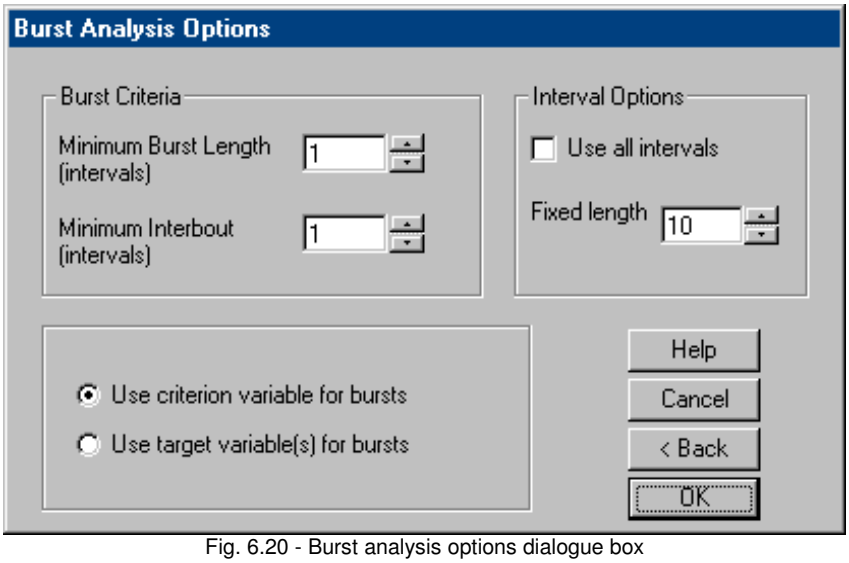


Fig. 6.20 - Burst analysis options dialogue box

Burst Criteria

A **Minimum Burst Length** of 1 interval will mean that all bouts of the criterion can be used in the analysis. If a minimum burst length of 10 was selected, then the bursts used will be restricted to those of the minimum length (10) and greater, ignoring those which are under 10 intervals.
The **Minimum Interbout** will join all bouts where the gap between the bursts is less than or equal to the figure selected. This allows successive occurrences of a variable to be grouped together (into longer bursts). This is equivalent to creating a new burst variable (see section 4.5).

Interval Options

The default is to use a fixed number of 10 intervals for each interval prior to, during and following each burst, although the user can opt to use all intervals.

The output from an interval burst analysis is shown in figure 6.21 below.

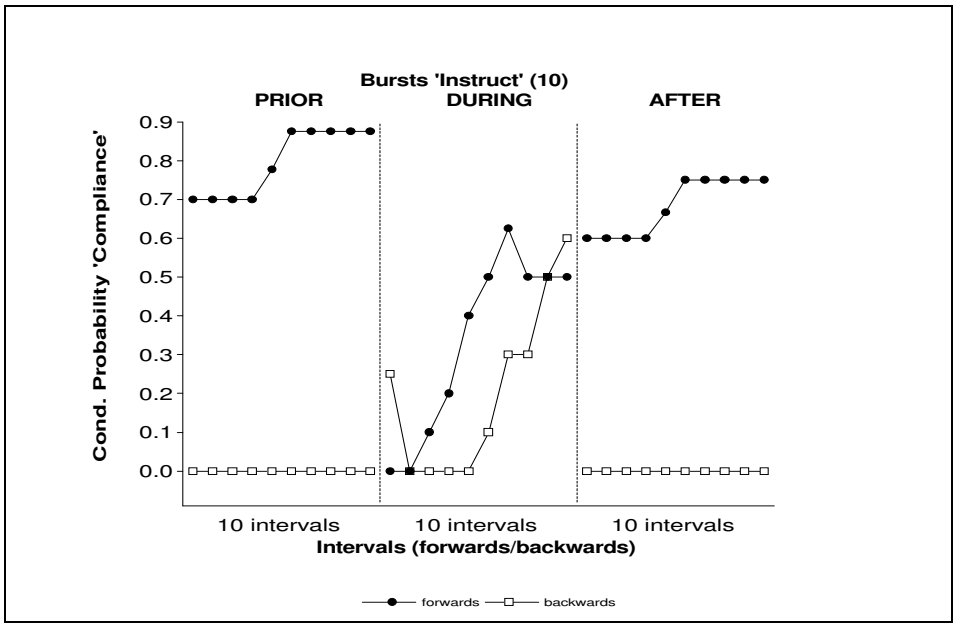


Fig. 6.21 - Burst analysis (using intervals)

(b) Burst Analysis (Using Percentiles)

This option offers a solution to the problem of differing burst lengths and interbouts, in that all the bursts and interbouts, regardless of their length, are divided into percentiles, and the conditional probabilities of the occurrence of other variables are then calculated for each percentile (see Hall and Oliver, 1997). This effectively scales the length of every burst proportionally.

For a percentile analysis choose

Analyze
→ **Burst Analysis**
→ **Using Percentiles…**

and select a criterion and one or more target variables.

The same options dialogue box will appear (see figure 6.20), although the **Interval Options** will not be present.

The output from a percentile burst analysis is shown in figure 6.22 below.

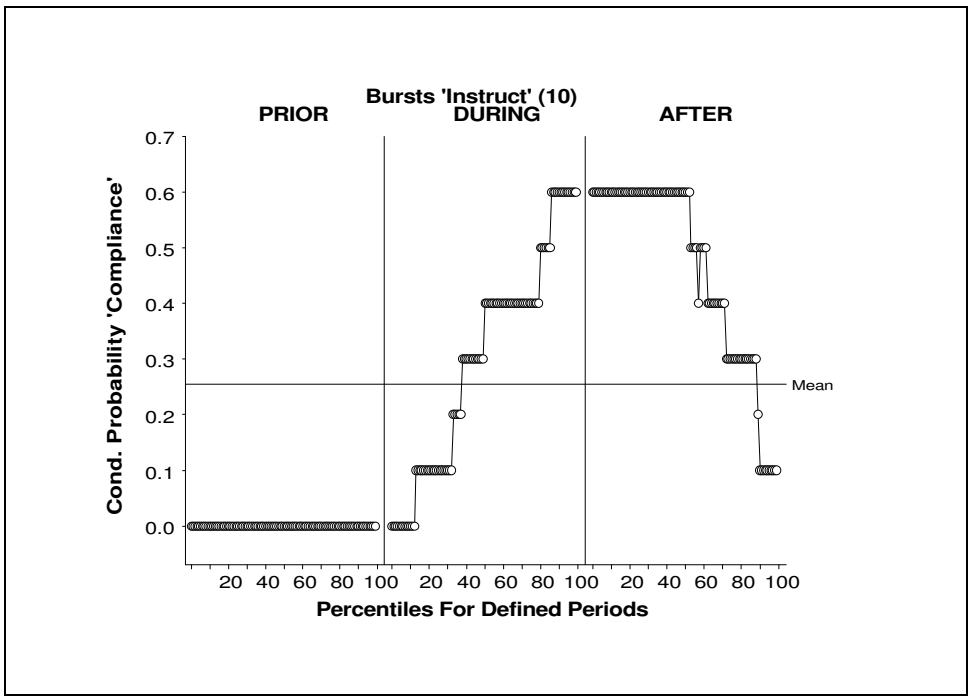


Fig. 6.22 - Burst analysis (using percentiles)

Note A number of statistics are included in the listing file for a percentile burst analysis. These include the local rate statistic (the sum of the number of intervals during which the criterion variable occurs within bursts divided by the total number of intervals during which the bursts occurred) and profile scores for each section of the graph (see Hall and Oliver, in press).

6.8 Survival Analysis

Survival analysis attempts to determine the time beyond which a variable is statistically unlikely to re-occur - this is also referred to as the *bout criterion***.**

Choose

Analyze
→ **Survival Analysis…**

and select one or more variables for the analysis.

A survival graph is created by plotting a cumulative distribution of interbouts against time. The output is in the form of a graphical display (see figure 6.23 below).

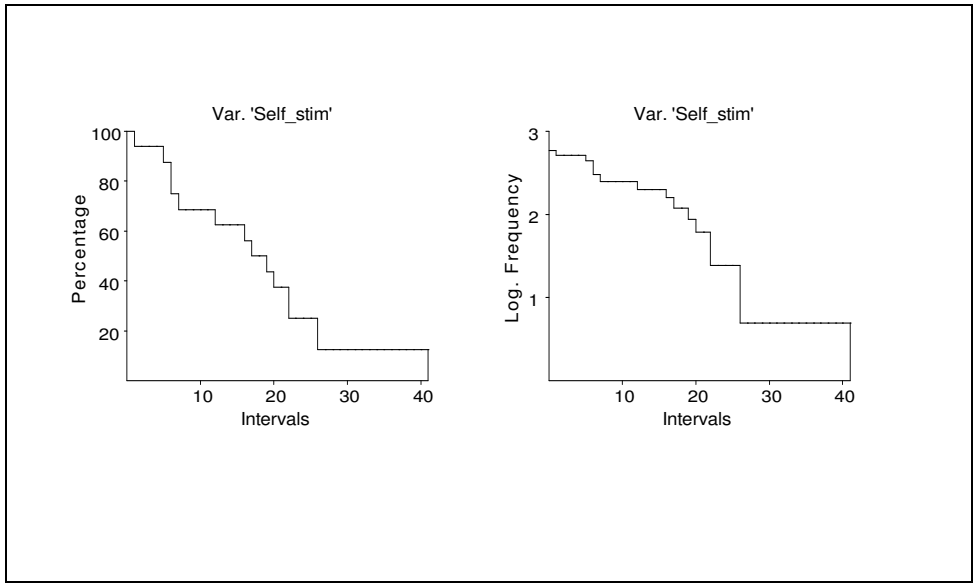


Fig. 6.23 - Survival analysis

Percentage Surviving

The slope of this graph will be a measure of the probability of a variable occurring at any time - a steep slope indicating a high probability of a further occurrence, a gradual slope indicating a low probability.

Graphs that are concave in shape (i.e. descending rapidly at first indicating that further occurrences are quite likely, then gradually approaching zero indicating that further occurrences are less likely) can be used to estimate the bout criterion for a variable.

Log Frequency Surviving

The log of the frequency surviving can be plotted and compared to other mathematical distributions - a straight line indicating random distribution and no change in the probability of occurrence over time.

Note Clicking on the  button on the main toolbar or pressing **F8** will repeat the last chosen analysis option using the same previously selected variables.

6.9 Analysis Options

Choosing

Analyze
→ Analysis Options

or clicking on the  button on the main toolbar allows the user to set various options such as creating filter and temporary (run-time) variables to determine how the dataset is analyzed.

Figure 6.24 shows the analysis options dialogue box.

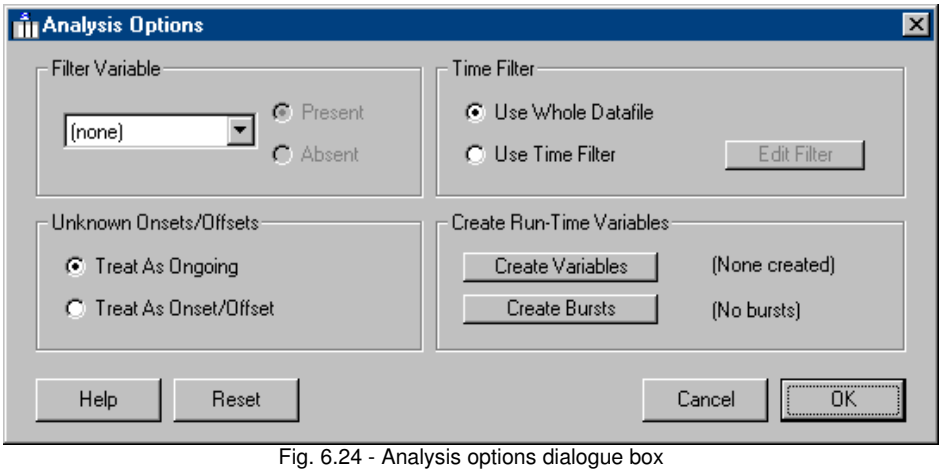


Fig. 6.24 - Analysis options dialogue box

Filter Variable

One may select a filter variable from the drop down list box (top left). A filter variable, present or absent, will restrict all analyses to those portions of each datafile where the variable is occurring or not occurring. The default setting is to have no filter variable.

Time Filter

Using a time filter will restrict all analyses to the period between the start and finish time intervals, which may be set by clicking the **Edit Filter** button.
With filesets, the time filter will apply to the same time periods in each of the files.

Note If the finish time exceeds the length of the datafile then all the data left will be used. If the start time exceeds the length of the datafile then there will be nothing to analyze.

Unknown Onsets/Offsets

Unknown onsets or offsets (i.e. variables that are already occurring at the start or end of a file) may be treated either as ongoing, or as starting/finishing in time with the beginning and end of the file. This will effect some of the analysis procedures (summary statistics, frequency distributions, and burst analysis).

Create Run-Time (temporary) Variables

Clicking on the **Create Variables** button, or **Create Bursts** button (bottom right), allows the creation of new composite variables (as described in section 4.5 - **Creating new variables**). These are temporary (existing only until the analysis options are changed or the datafile/fileset is closed - permanent variables should be created from the main **Edit** menu option).
Click on the button and follow the prompts (refer to section 4.5 for guidelines).

Note Once set, all the analysis options will apply to all forms of analysis (under the **Analyze** main menu), apart from the Inter-Observer Reliability. If temporary variables have been created they can be used for the various graph options.

---

7. Inter-Observer Agreement (IOA)

Inter-observer agreement measures inter-rater reliability for any number of variables.

Choose

Analyze
→ Inter-Observer Agreement…

or use the **F9** key. The following dialogue will appear:

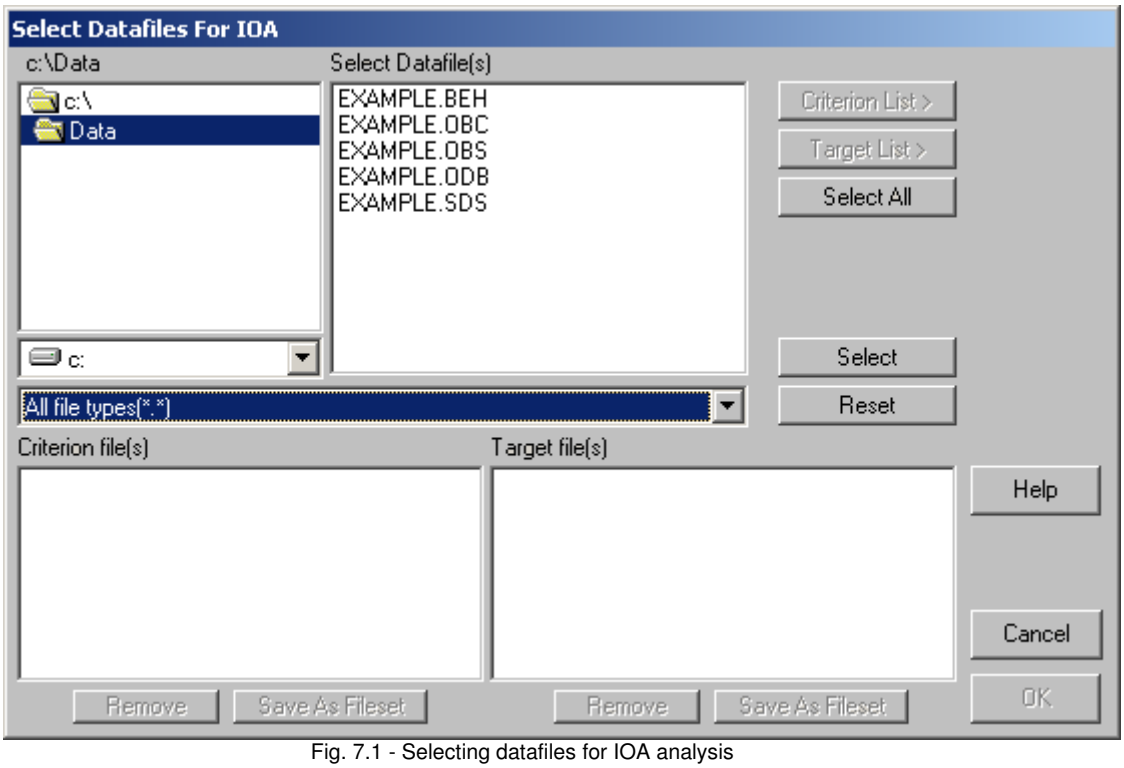


Fig. 7.1 - Selecting datafiles for IOA analysis

Select a datafile in the **Select Datafile(s)** list box and use either the **Criterion List >** button or the **Target List >** button to add the datafile to the respective list.

Note Criterion list files are used for primary observer datafiles - secondary observer datafiles are called target datafiles.

Datafiles can be deselected from the criterion or target list by selecting them and pressing **Remove**. Selection criteria options are present (see **File Selection Criteria** in section 3.4).

The criterion or target list can be saved as a fileset (providing there is more than one file in the list).

IOA Options

Click on **OK**, and the **IOA Options** dialogue will appear:

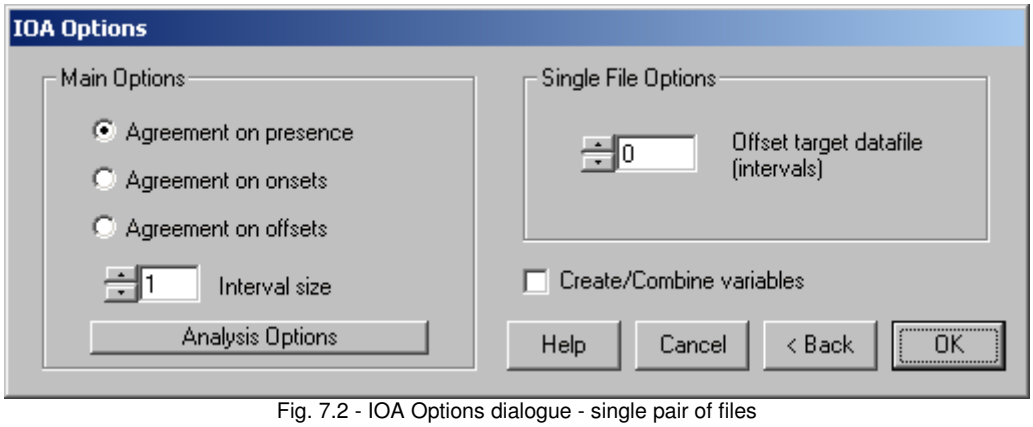


Fig. 7.2 - IOA Options dialogue - single pair of files

The user can control how stringent the analysis is with several options.

Main Options

Selecting **Agreement on presence** (the default) allows all the intervals in which a variable occurs to be used in the analysis. Selecting **Agreement on onsets** restricts the analysis to the onsets of specified variables. Selecting **Agreement on offsets** restricts the analysis to the offsets of specified variables.

The different analysis options are illustrated in figure 7.3.

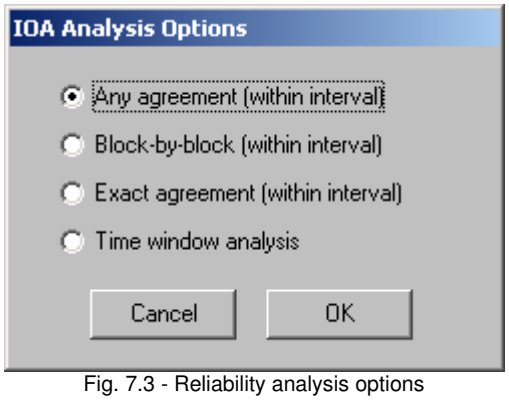


Fig. 7.3 - Reliability analysis options

Any agreement (within interval) checks to see whether two observers agree that the occurrence of a specific variable has occurred within the interval specified by **Interval Size.** This option performs a temporary Partial Interval Conversion on each dataset. The data are broken up into intervals (e.g. if an interval of 10 is used then the intervals are 1 - 10, 11 - 20, 21 - 30 etc.), and the occurrence or non-occurrence of a specific variable in matched intervals are compared between datafiles. Agreements are calculated on an interval by interval basis and presented as total percentage agreement (R tot), percentage occurrence agreement (R occ), percentage non-occurrence agreement (R nonocc) and (Cohen's) Kappa.
When percentage occurrence agreement, percentage non-occurrence agreement and Kappa values cannot be calculated (i.e. when there is no occurrence of a variable or it occurs in every interval) the data are displayed as "**-.--**".
The number of file pairs used in the calculation of Kappa (if using more than one pair) is also indicated at the right of the display grid.

In the **Block-by-block (within interval)** method, the smaller of the two observers' totals (number of seconds of occurrence of a specific variable) in an interval (specified by **Interval Size**) is divided by the larger. This provides a score of between zero and 1.0 for every interval. Where both observers scored non-occurrences during an interval, such intervals are scored as 1.0 (perfect agreement although mathematically zero cannot be divided by zero). Scores are summed across all intervals, divided by the number of intervals, and the result multiplied by 100 to provide a percentage agreement index.

In the **Exact agreement (within interval)** method, exact agreement coefficients are calculated by partitioning each observers datafiles into intervals (specified by **Interval Size**). In each interval, two observers can agree on the exact number of seconds a specific variable occurred, agree that the variable did not occur, or disagree about the exact number of seconds that the variable occurred. Percentage agreement coefficients are calculated by dividing the number of agreements by the sum of agreements plus disagreements and multiplying by 100.

The **Time window analysis** method involves second-by-second comparisons across the two datafiles with **Tolerance** (specified by. ± t seconds) either side. When both records show the occurrence of a specific variable (± t seconds) this second is counted as an agreement. Any second (± t seconds) in which only one record contains an occurrence of the specific variable is a disagreement. Percentage agreement coefficients are calculated by dividing the number of agreements by the sum of agreements plus disagreements and multiplying by 100.

Note The most stringent analyses for each of the IOA Analysis Options (using an Interval Size of 1 second for the **Any Agreement**, **Block-by-block** or **Exact Agreement** methods or a Tolerance of ± 0 for the **Time window analysis** method) produces an equivalent analysis.

Note Irrespective of the type of analysis option selected, the resulting data display also shows agreement based on frequency and duration using the following formula: the lowest frequency/duration of a variable divided by the highest frequency/duration multiplied by 100.

Single File Options

If reliability is being calculated for one pair of files then the user can **select Offset target datafile** to ensure that the two files are correctly aligned. This will work if the target datafile is consistently *n* intervals in front of or behind the criterion datafile, which is possible if the timers were not started in synchrony.

Multiple File Options

If more than one pair of datafiles are being analyzed then the **IOA Options** dialogue box has a **Multiple File Options** section instead of a **Single File Options** section (see figure 7.4 below).

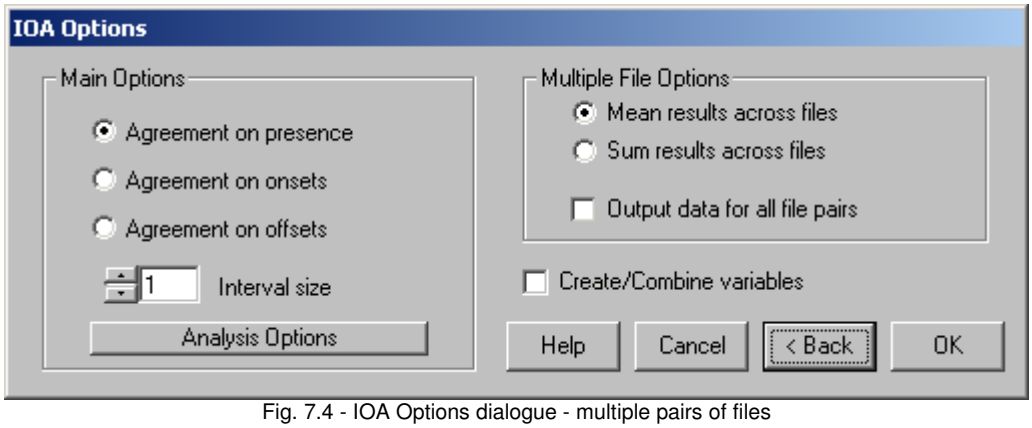


Fig. 7.4 - IOA Options dialogue - multiple pairs of files

If agreement is being calculated for more than one pair of files then the results can either be the **Mean results across files** (this can be useful where file lengths are all very similar) or one can **Sum results across files** and the overall reliability data presented. One may also **Output data for all file pairs** by checking this option - reliability data for each of the file pairs will then be output to the listing file only.

If the agreement data are averaged across multiple files, then the ranges are also shown in the results grid.

Create/Combine Variables

Checking this box allows the user to create temporary variables for use in the IOA analysis - (see **Creating new variables**, section 4.5).

Click on **OK** and a **Select Variables** dialogue box appears, allowing the user to check agreement for one or more variables.

Choosing

> **Agreement**
> → **Another Analysis**…

returns to the file selection dialogue (see figure 7.1).

Choosing

> **Agreement**
> → **Change Options**…

returns the user to the options dialogue (see figures 7.2 and 7.3).

Appendix

a) Compatibility (File Types/Conversions)

ObsWin will read the following data formats;

1. Timed Sequential Data (*.SDS datafiles - see Bakeman and Quera, 1995)
2. ObsWin comma delimited files (*.OBS datafiles)
3. ObsWin compressed files (*.OBC datafiles)
4. Communitech International files (*.BEH datafiles)
5. Psion 'CTS' files (*.ODB datafiles).
6. The Observer 3.0 files (*.ODF datafiles).

ObsWin creates and saves observational data text files in Timed Sequential Data format, ObsWin comma delimited format or ObsWin compressed format.

b) Portability

For portable data collection two different formats of the data collection program are supplied with the main program.

A DOS version of the data collection program (COLLECT.EXE) can be loaded and run from a notebook or palmtop PC (not all notebooks have the power or memory to run Windows). The same set of basic data collection options are available (see Section 2.3) although variable names are restricted to single uppercase single characters and cannot be relabelled from within the DOS program.

A Windows CE version of the data collection program (ObsWin.vb) can be loaded and run from a handheld (H)PC or palm-sized (P)PC. Again, variable names are restricted to single uppercase single characters and data may only be collected to an accuracy of 1-second.

Datafiles generated from the DOS version can be saved in either *.SDS or *.OBS format. Datafiles generated from the Windows CE version may only be saved in *.SDS format.

c) Technical Support

Contact:

ObsWin
c/o Antam Ltd.
3 Dallin Road
Shooters Hill
London SE18 3NY

Email:
Neil@antam.co.uk

Web:
http://www.antam.co.uk/obswin.htm

To address a specific problem or query e.g. a potential bug with the program, then please send a complete description of the problem along with a copy of the datafile(s) pertaining to the problem (if relevant). Any bugs found will be rectified as quickly as possible.

References and suggested reading

Bakeman,R. & Gottman,J.M. (1997), *Observing interaction: an introduction to sequential analysis* – second edition, Cambridge: Cambridge University Press.

Bakeman,R. & Quera,V. (1995), *Analyzing interaction: sequential analysis using SDIS and GSEQ*, Cambridge: Cambridge University Press.

Emerson,E., Reeves,D., Thompson,S., Henderson,D., Robertson,J. & Howard,D. (1996) Time-based lag sequential analysis and the functional assessment of challenging behaviour, *Journal of Intellectual Disability Research, **40***, 260-274.

Gottman,J.M. & Notarius,C. (1982), Sequential analysis of observational data using Markov Chains, In *Strategies to Evaluate Change In Single Case Research*, Kratochwill,T (Ed.), Academic Press, New York.

Greenhouse,J.B., Stangl,D. & Bromberg,J. (1989), An introduction to survival analysis: statistical methods for analysis of clinical trial data, *Journal of Consulting and Clinical Psychology*, ***57**,* 536-544.

Hall,S. & Oliver,C. (1997), A graphical method to aid the sequential analysis of observational data, *Behavior Research Methods, Instruments & Computers, **29**,* 563-573.

Harrop,A., Daniels,M. & Foulkes,C. (1990), The Use Of Momentary Time Sampling And Partial Interval Recording In Behavioural Research, *Behavioural Psychotherapy*, ***18***, 121-127.

MacLean,W.E., Tapp,J.T. & Johnson,W.L. (1985), Alternate methods and software for calculating interobserver agreement for continuous observation data*, Journal of Psychopathology and Behavioral Assessment*, ***7***, 65-73.

Mansell,J. (1985), Time Sampling And Measurement Error: The Effect Of Interval Length And Sampling Pattern, *Journal Of Behaviour Therapy And Experimental Psychiatry, **16***, 245-251.

McGill,P., Hewson,S. & Emerson,E. (1994), *CTS - A Real Time Event Recorder for the Psion Organiser*, Canterbury: Tizard Centre, University of Kent.

Murphy,G. (1987), Direct Observation As An Assessment Tool In Functional Analysis And treatment, In Hogg,J and Raynes,N,V (Eds.) *Assessment In Mental Handicap: A Guide to assessment practices, test and checklists*, Croom Helm, London.

Noldus Information Technology (1995), The Observer - Version 3.0, Noldus: Wageningen.

Repp,A., Harman,M.L., Felce,D., Van Acker,R. & Karsh,K.G. (1989), Conducting behavioural assessments on computer-collected data, *Behavioural assessment*, ***11***, 249 - 268.

Repp,A.C. & Felce,D. (1990), A Microcomputer System Used For Evaluative And Experimental Behavioural Research In mental Handicap, *Mental Handicap Research*, ***3**,* 21-32.

Sackett,G.P. (1987), Analysis of sequential social interaction data: some issues, recent developments, and a causal inference model. *Handbook of Infant Development* , 2nd Ed.

Taylor,J.C. & Romanczyk,R.G. (1994), Generating hypotheses about the function of student behaviour by observing teacher behaviour, *Journal of Applied Behavior Analysis, **27***, 251-265.

Index