

DATA ENTRY GUIDELINES

VANUATU DHS 2013

1. GENERAL GUIDELINES FOR DATA ENTRY

The data entry for the DHS survey will be carried out using the software package CSPro. CSPro is an interactive data entry system that can check for acceptable codes for questions, follow skips and filters in the questionnaire and check the consistency of data as they are entered. The DHS questionnaires will be entered by cluster, with each cluster being assigned to one data entry operator. The data for each cluster will be entered into a separate data file for that cluster on the hard disk of the computer, rather than into one large data file. This is to protect against a major loss of data due to hardware or software failure.

During data entry it is important to remember that speed of data entry is not important but accuracy of entry is important. CSPro is like a computer-aided office editor. It will not allow any values to be entered that it thinks are out of range. It will skip to the next question that it believes is the next to be entered. It will present a new screen when it thinks a new screen is necessary. CSPro does these things according to the data entered. It can not check everything that the operator enters. If an operator enters a value that was not recorded on the questionnaire, but is a value that is within range, CSPro will go to the next field that this value demands. Each operator must review what they have entered as they are entering data to ensure that mistakes do not take place.

Because CSPro will control which question is the next to be entered, it is important for the operator to watch the screen, to see which field CSPro expects to be entered next. Operators should not enter data with their heads down, only reading the questionnaire. This will lead to many errors in the data entered if the operator miskeys any data or if there are any mistakes in the responses recorded on the questionnaire. At a minimum, the operator should look at the screen every time a page is turned in the questionnaire.

Data are recorded on the questionnaire in two ways: (1) codes are circled, or (2) numbers are written in boxes. Some questions allow the response to be in either of the two forms or in a combination of the two forms. For example, «Q214» uses circled codes, «Q101» uses codes written in boxes, «Q102» uses either codes written in boxes or circled codes, and «Q406» uses a circled code followed by a code written in boxes.

If during data entry the operator finds that there are data in the questionnaire which she/he has not been asked by CSPro to enter or she/he finds that CSPro is demanding an answer to a question that was not asked in the survey, then the operator should review all of the data entered to ensure that she/he has not entered the wrong value in one of the questions. When reviewing the data entered check question by question whether the data entered is the same as the response written on the questionnaire. Do not look at just the last few questions, but look back over the whole page and the preceding page to ensure that no typing error was made in an earlier question.

If there was no mistake by the operator in entering data then she/he should ask the data entry supervisor to resolve the mistake. If the supervisor is unable to resolve the mistake then the problem should be shown to the senior survey staff for resolution.

During data entry, messages may be produced which will require attention and correction to the data entered. Operators should read the message carefully before attempting to resolve the problem. Resolve the problem in the same way as before, reviewing the data entered first to check that no entry error was made. If no entry error occurred then ask the supervisor to resolve the problem. The supervisor will have a manual containing a list of messages and ways to resolve the problems.

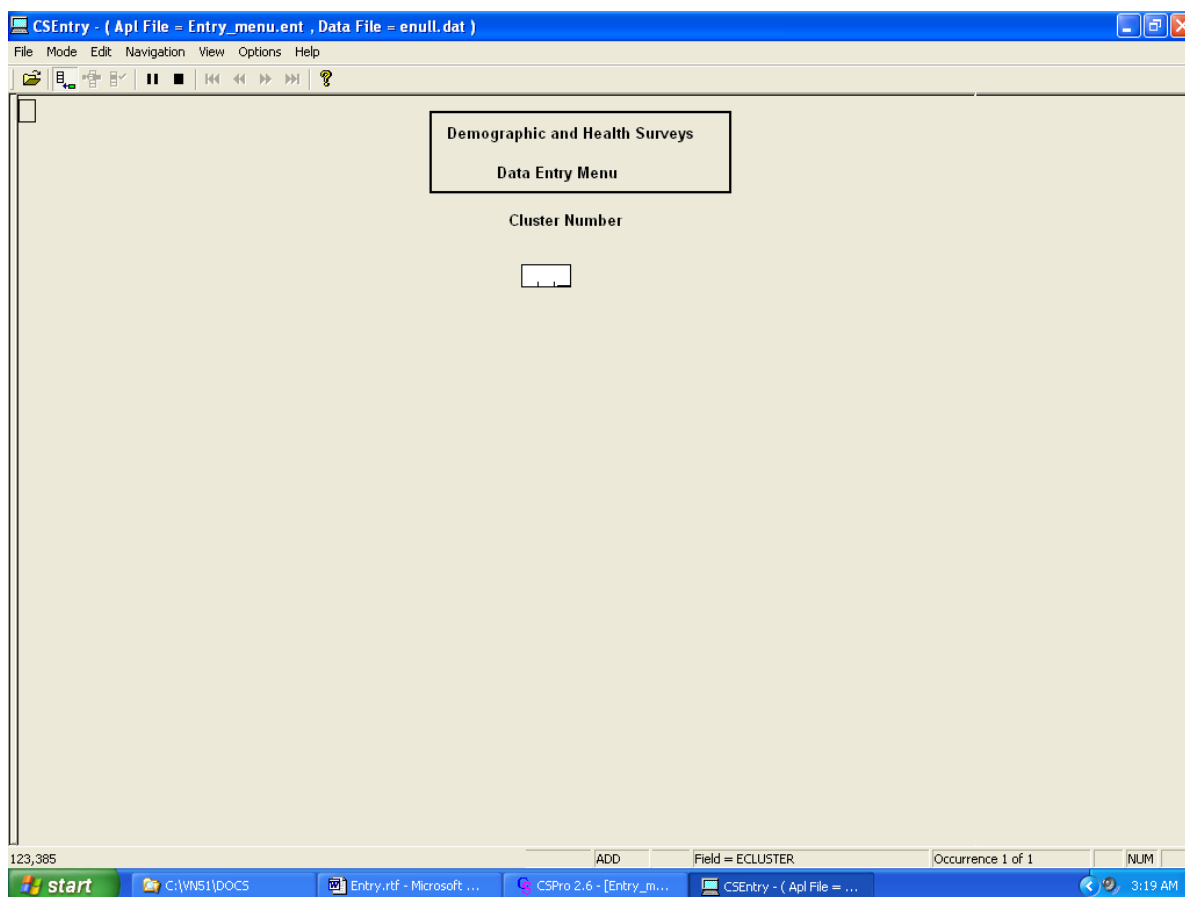
During data entry keep all questionnaires in a tidy pile. Once a questionnaire has been entered turn the questionnaire over onto a second pile. The questionnaires should be kept in numeric order by household at all times, and within households individual questionnaires should be kept in order according to line number of the respondent. At the end of data entry for a cluster re-tie the pile of questionnaires with the cluster control sheet on top and pass the cluster to the supervisor, who will then assign a new cluster for entry.

2. STARTING DATA ENTRY

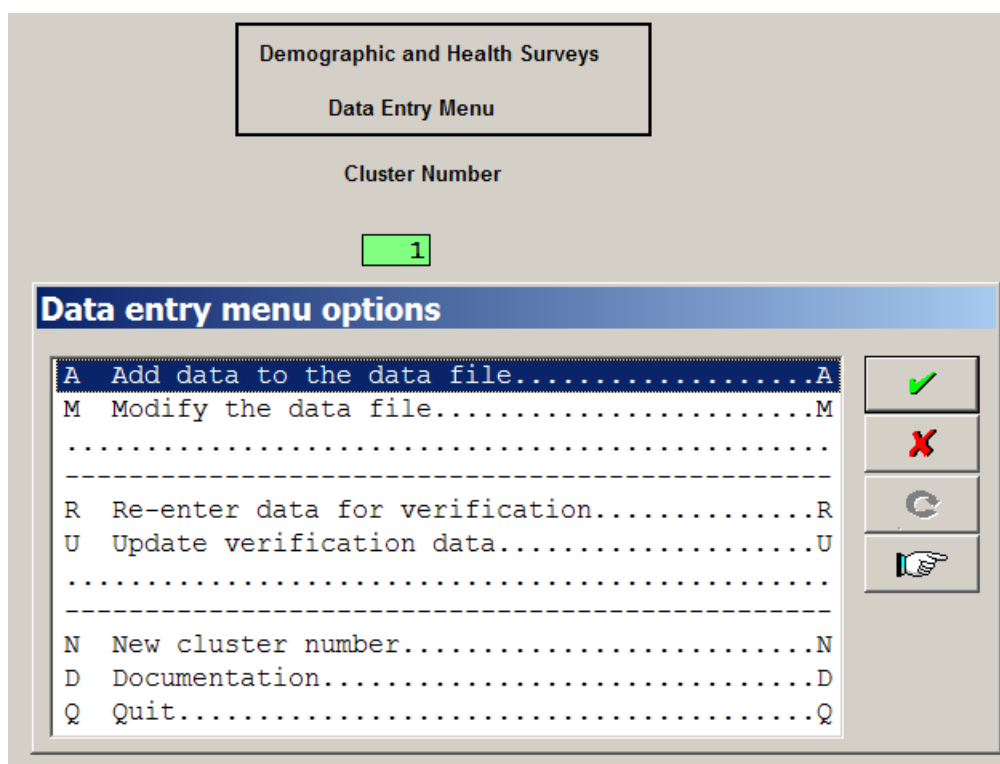
a) Data Entry Menus

To start CSPro data entry menu:

- Make sure the computer date is correct.
- Click the DHS Data Entry icon on the Windows desktop. At this point all files required by the data entry program will be transferred from the supervisor to the data entry computer. It also transfers the control and the sample files. This is necessary to make sure that the data entry machine has the latest versions of the applications. This is particularly important at early stages of the data entry operation as the application needs to be tuned up.
- The following screen will appear, requesting a cluster number be entered. The data entry menu uses the control file to determine if the cluster is valid. If the cluster is not part of the survey sample, the program will not let the data entry operator to proceed.



1. After typing a valid cluster number, the Data Entry menu screen below will appear. To make sure that the cluster entered is valid, the data entry menu uses the control file.



The menu shown above gives various options:

A Add data to the data file. This option is used to enter data for the main data file. The data is stored in the \DATA directory of the data entry machine. After choosing this option the menu system will check if the cluster has already been assigned for main entry by the data processing supervisor.

M Modify the data file. This option allows corrections to be made to questionnaires already entered into the "Main" data file. The option will be valid only if it has already assigned for main data entry and there is already a data file for that cluster in directory \DATA.

R Re-enter data for verification. This option is used to enter data that are being verified or entered for the second time. It is identical to option A, except that a different data file is being used for the questionnaires. The data is stored in the \VERI directory of the data entry machine. After choosing this option the menu system will check if the cluster has already been assigned for verification by the data processing supervisor.

U Update verification data. This option is used to modify the verification data entered with option R. It is identical to option M, except that it uses the verification data file. The option will be valid only if it has already assigned for verification and there is already a data file for that cluster in directory \VERI.

N New cluster number. This option will be used to change or correct the cluster number that is currently being processed. When data entry for a cluster has been completed, this option is used to change to a new cluster number.

D Documentation. This option will display this documentation on the screen and allow the user to browse through the documentation.

After exiting the data entry application either from main entry or verification, the menu system will transfer the data from the data entry to the supervisor machine. If the data entry operator was entering data to the main data file the cluster file will be copied from directory \DATA in the data entry computer to directory \DATA in the supervisor machine. If the data entry operator was verifying data the cluster file will be copied from directory \VERI in the data entry computer to directory \VERI in the supervisor machine. These copies are made for either add or modify mode.

If you are entering data into the "Main" data file, press option A; if you are verifying data, press option R. The first data entry screen will be displayed.

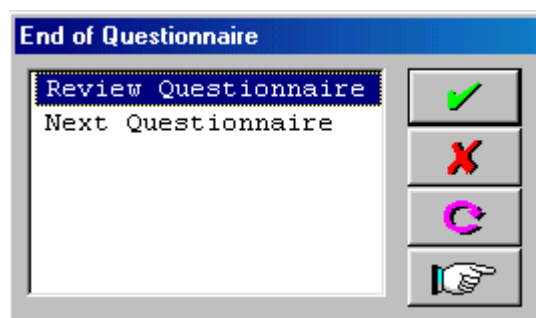
Press the Num Lock key by the numeric keypad to the right of the keyboard so that the Num Lock light is shining (if a light exists). This will permit the use of the numeric keypad for the entry of data. The numeric keypad should be used for all numeric data entry. The alphabetic keys will only be used to enter textual responses specified against the category "Other" and for the questions with alphabetic codes.

During data entry, enter the values shown on the questionnaire including all leading zeros.

b) Household Schedule Cover Sheet

Each screen is laid out in a similar manner to the questionnaire. On the first screen enter the «QHISLAND» Island, «QHVILLAGE» Village, «QHEA» Enumeration area, «QHNUMBER» Household Number, «QHURBRUR» Urban/Rural, and «QHELIGM» Male Survey Selected codes in order. The program will automatically go from one field to the next field when a field is full; it is not necessary to press ENTER after entering each field. After the Cluster Number and Household Number codes have been entered the program will check to see if this household has already been entered. After this, continue with the «QHINTD» Day, «QHINTM» Month and «QHINTY» Year of Interview, «QHINTNUM» Interviewer Number, and «QHRESULT» Result Code, «QHVISITS» Total Number of Visits. If the Result Code is 1 (Completed) then enter the «QHINTLAN» Language of interview, «QHRESLAN» Language of respondent, «QHTRANS» Translator used, «QHMEMBER» Total in Household, «QHWOMEN» Total Eligible women, «QHMEN» Total Eligible Men (if applicable), and «QHRESP» Line Number of the Household Respondent. If the Result Code is not 1 then the program will skip these fields. Finally enter the «QHSUPERV» Supervisor code, «QHFEDIT» Field Editor code, «QHOEDIT» Office Editor code and the «QHKEYER» Keyer code. The Keyer code is the operator code assigned to each operator. If you are entering data into the "Main" data file, your keyer code should be written on each questionnaire. If you are verifying data, only enter your keyer code but it should not be written on the questionnaire.

If the result code was not 1 then that is all that will be entered for that questionnaire and CSPro will display the following:



First check that there are no more data to enter for the questionnaire before continuing. By typing the letter R (Review Questionnaire) the data entered can be reviewed and corrections made, if necessary. If the letter N (Next Questionnaire) is selected, CSPro will then show a new household cover sheet for the next household questionnaire. Only continue with the next questionnaire after checking each of the fields to ensure that they are correct.

If the result code was 1 then CSPro will show the next screen for the household members and give the line numbers of each of the household members that it is expecting to be entered.

c) Household Schedule

Check that the number of lines with line numbers on the screen is the same as the number of lines completed on the questionnaire.

Enter each member in order, entering data across the lines. When the eligibility columns (columns 9,10, 11a, 11b, 11c and 25) are reached enter the number from that column if it has been circled, but otherwise enter 00 (except in column 25 that always contains the line number) if it has not been circled (CSPro will automatically put 00 in for persons who are not eligible in most cases). Continue entering the members until all members have been entered. After entering the last data for the last member, CSPro will present a similar screen to the one shown above. Check all of the entries on the page to ensure that they have been entered correctly, and then select the letter R to review or correct data on the schedule, or the letter N to go to the next section

After completing the household schedule, CSPro will present the screens to capture child labor, child discipline and disability modules. In child labor it's necessary to input all children between 5 and 14 years old (those children are marked in column 11b in the household schedule). In the child discipline module only 1 selected child will be input and in the disability module it's necessary to input all persons with disability (those are marked in column 25 in the household schedule)

After completing those sections, CSPro will present the screen for the mosquito nets, the household characteristics, and the screens for the height and weight of children, women and men. If there are no eligible children, women or men, the respective screen will not be displayed.

After the height and weight sections, we are ready to enter the first individual (woman or man) questionnaire if there are eligible respondents.

If there are no eligible women and no eligible men in the household then CSPro will present the screen for the next household questionnaire, but if there are eligible women or eligible men, CSPro will display the screen for the Individual Questionnaire cover sheet. If the Household Schedule cover sheet is shown instead of the Individual Questionnaire cover sheet and there are individual questionnaires to be entered, ask the supervisor for assistance in correcting the problem.

d) Individual Questionnaire Cover Sheet

When the Individual Questionnaire cover sheet is displayed the requested information needs to be filled in as similar way to the household questionnaire. Check that each of the identification variables are correct before continuing, and then enter the «QLINE» Line Number of the Woman or Man. **The individual questionnaires must be entered in order according to that line number.** Next enter the remaining fields on the Individual Questionnaire Cover Sheet.

If the result code was not 1 then that is all that will be entered for that individual questionnaire and CSPro will again display a screen, asking for the data to be reviewed or to continue with the next questionnaire. Only continue after checking that each entry on the page is correctly entered.

e) Individual Questionnaire

Continue entering data in the same manner for each of the questions in the questionnaire, verifying that the correct route is being followed at all times and that no fields are being skipped that have data in the questionnaire. Throughout the entry of the questionnaire, ensure that the data on the questionnaire is the same as the data entered. Do not try to enter the data too quickly. It is important that the data are correct and speed of data entry is not crucial.

There are certain key sections of the questionnaire that will require special attention and these are explained in further detail in section 8.

At the very end of the questionnaire, CSPro will again display a screen allowing the data entered to be reviewed or to continue with the next questionnaire. Only continue after checking that the data entered are correct.

After an individual questionnaire is complete, CSPro will display the cover sheet screen for an individual questionnaire if there are more individuals to enter for the household, or the cover sheet for the household if the individual just entered was the last individual in the household.

3. MOVING BACK THROUGH THE DATA

If a mistake is made during data entry then it is possible to go back and correct the mistake provided the mistake is found before the end of the questionnaire. It is possible to go back to questions on previous screens as well as to questions on the current screen. There are two ways of going back to earlier questions:

- a) Pressing up arrow or Shift Tab key to go back one field. Repeating this will go back another field. This can be used to move back to the field where the correction is to be made.
- b) Pressing the function key F6 (GoTo) will ask for the name of a field to go back to. Enter the name of the variable to correct and then click OK to go to that field. If the field is part of a repeating section of the questionnaire then an occurrence number can be given by pressing the Tab key after giving the name of the field, entering the occurrence number and then clicking OK.

In the individual questionnaire the variable names are based on the question numbers preceded by the letter Q for woman questionnaire and QM for man questionnaire, and sometimes followed by another letter if there is more than one field for the question. Generally the letter after the question number will follow the pattern A, B, C, D... although there are some exceptions. If the question asks for months and years then the two parts are suffixed M and Y. If the question asks for hours and minutes then the two parts are suffixed H and M. If the question asks for units and numbers then the two parts are suffixed U and N. Some examples are «Q101H», «Q101M», «Q106M» and «Q106Y». Names of variables in the household questionnaire cover sheet and on the individual questionnaire cover sheet are given by name and the list of variables should be used to refer to these variables. Variables in the household schedule and the other sections of the household questionnaire consist of QH followed by the number of the question, e.g.

«QH03»or «QH10». A list of all variable names is available with the data entry supervisor.

4. CORRECTING DATA

If a mistake has been made in the middle of entering data into a numeric field, pressing the Space bar can erase the information already entered in that field. For example, if code 01 was to be entered for «Q217» but the 1 was pressed first then the whole field can be erased by pressing the SPACE bar. The corrected data can then be entered into the field. The Backspace and Del keys cannot be used to correct data entered in numeric fields.

If a previously entered numeric field contains a mistake, the correct data can be entered over the top of the previously entered data. However, if the operator is not sure, she/he should use the SPACE bar to empty the field before entering the data.

In correcting data in alphabetic fields, the Backspace, Del and Insert keys can be used. The Space bar may be used to enter a single blank character in these fields, but does not completely blank out the field.

5. MOVING FORWARD THROUGH THE DATA

Having moved back through the data to an earlier field there are two ways of moving forward again.

- a) Press the down arrow key or Tab key for each field until the required field has been reached.
- b) Pressing the function key F6 (GoTo) will request the name of a variable to advance to. Giving the name of the variable and then clicking OK, as for moving back through the data, will move to the variable requested. If it is desired to move to the last variable entered in the questionnaire, press F10 (Advance to end) without specifying the name of a variable. It is not possible to move forward further than the last variable entered, so CSPro will advance only as far as this variable.

6. SKIP ERRORS AND MISSING DATA

CSPro is programmed to present on the screen each of the questions that should be keyed by the operator. As the response for each question is entered, the operator creates a data path through the questionnaire. Certain responses are used to help control this path. For example, prior to entering the household schedule, the data entry operator will be asked to enter the total number of persons in the household. This field is then used to control the entry of the household schedule. Similarly, the total number of eligible women (and/or men if included) is used to control the number of individual questionnaires entered for the current household.

In the individual questionnaire there are also responses to questions that control the data path. For example the year of birth of each child in the birth history

is used to control whether the child should be included in the health section of the questionnaire. Any field that determines a skip is also controlling the data path.

If the fields that control the data path are miskeyed, the data entry operator may be asked to enter responses that do not exist on the questionnaire or may not be allowed to enter responses that do exist in the questionnaire. This is because the data entry operator is following an incorrect data path. If this occurs, the operator should return to the "control" field that determined the skip and correct the entry for that field.

Occasionally a skip error will exist in a questionnaire. If a woman has reported YES according to question «Q216», but the following question records no answer, the interviewer has committed a skip error. However, an operator should never assume that the previous keying has all been correct and that the error is in the questionnaire. The operator should check that the code recorded in question «Q216» has been properly entered.

In general, if a skip error is found in the questionnaire and data are missing the operator records a special code indicating missing data. For numeric fields, this code is a field full of 9s. For alphabetic fields, this is a single question mark "?" recorded in the first position of the field. Any corrections made during data entry should be written into the questionnaire with a green pen.

It is a good habit to review other responses prior to the field that is identified as giving an error. Occasionally a data entry operator can be "off" by one field and enter many erroneous responses without the errors being detected. This can often happen in sections of the questionnaire that contain several Yes/No questions in sequence, e.g. «Q901» to «Q909». It is essential for operators to review earlier sections of the questionnaire when an incorrect data path appears to have been followed. If an operator is unsure of the problem, she/he should never hesitate to call a supervisor for help.

Operators should make certain that they enter the responses found in the questionnaire and not create responses in order to satisfy CSPro's demand for a response. A supervisor should always be consulted for those cases where serious skip errors are found in a questionnaire. These errors should also be recorded on the Problem Log kept by each operator and supervisor.

7. DATA PATH

CSPro remembers the current data path at all times and shows all values that are on the current path in green. When moving back through the data CSPro will change the fields after the current field to white. If a correction to an earlier field changes the flow of questions and a field that had data in it is bypassed then this field will become gray. At the end of entry of a questionnaire only those fields that are green are written to the data file, any fields that are not on the correct path are not written to the data file.

The reason for explaining this is that if a mistake is made such that data are entered on the wrong path then it is not necessary to go to those fields and blank out

the values entered in them, providing that the final path through the questionnaire is correct.

8. QUESTION DETAILS

a) Alphabetic code fields

There are several fields that have alphabetic codes for the responses and these are either circled or not. Enter all of the codes that are circled for the question and then press the ENTER key. The codes must be entered in order, with no gaps between the codes and no repetitions of the codes (for example, «Q334», «Q407B», «Q408», etc.).

b) Other answers

Each "Other" answer recorded in the questionnaire is to be recorded. When the screen for the other answer is displayed, type in the text of the response exactly as it is given on the questionnaire. Do not abbreviate the text; give as much detail as possible. After entering the text of the response, press enter to return to the data entry screen. The other answers text don't need to be enter in the secondary entry (verification).

9. ENDING QUESTIONNAIRES

When coming to the end of data entry for a questionnaire, review the data entered to ensure that no mistakes have been made. After finishing the entry of a questionnaire there is no chance of correcting the data on that questionnaire until later in a "Modify data" session. If a mistake was made but not discovered until after the end of the questionnaire, record the cluster number, household number, individual's line number, question number (and occurrence number if part of a set of repeating fields), with the error and ask the data entry supervisor to assist you in correcting the problem.

After the end of a questionnaire the data for that questionnaire is not saved until the operator clicks Yes when asked to Accept the questionnaire. This will bring up the first screen for the next questionnaire, i.e. a new household questionnaire screen or a new individual questionnaire screen.

10. STARTING A NEW CLUSTER

When a cluster of questionnaires has been finished, a new cluster will be assigned. Each cluster is to be entered in a different data file. The cluster number to use must be changed. Leave the data entry program by pressing the function key Alt+F4 (File, Exit). This should only be done when the computer is displaying the household questionnaire cover sheet and awaiting entry in the first field of the screen. After ending the entry program, the data entry menu will then be displayed. Use option N to change the cluster number to the new cluster number. Press option A to commence data entry for that cluster if entering data for the first time in the "Main" data file or press option R to reenter data into the "Verification" data file.

11. QUITTING DATA ENTRY

The only point at which it is possible to quit data entry is at the very beginning of a new household schedule cover sheet. Pressing the function key Alt+F4 (Exit CSEntry) at this point will leave the data entry program and return to the data entry menu. Pressing Esc will quit the menu completely and return to Windows desktop.

Never leave the computer in the middle of data entry. If the operator wishes to leave the computer then she/he must finish the household that she/he is entering and then press Alt-F4 to end and return to the data entry menu. **Always leave the computer on the data entry menu.**

As the end of the data entry session or the end of the day approaches, remember the amount of time it takes to enter a household and all of its individual questionnaires. Do not start a new household unless there is time to properly complete the household.

----- End of Document----