

Main Study

Manual for Improving Data Quality

SACMEQ III
The Southern and
Eastern Africa
Consortium
for Monitoring
Educational Quality

September 2007

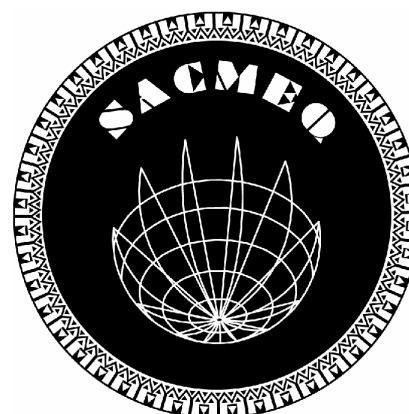


Table of Content

Content	Page #
Introduction	5
Section 1: Special Instructions on Data Quality Check <u>during Data Collection</u>	7
Section 2: Special Instructions on Data Editing <u>before Data Entry</u>	17
Section 3: Special Coding Instructions <u>during Data Entry</u>	33
Section 4: Special Instructions on Data Verification <u>after Data Entry</u>	41

Introduction

This document has been prepared in order to provide you with some tips and guidance for achieving high quality data sets for SACMEQ III.

The contents have been divided into four parts: (i) during data collection, (ii) before data entry, (iii) during data entry, and (iv) after data entry.

The National Research Coordinators (NRCs) should read all of this material, make several copies, and then distribute different sections of the document to the relevant people.

Section 1: Special Instructions on Data Quality Check during Data Collection

<p>This section has been already included as Appendix B in the Manual for the Data Collectors</p>

The Data Collectors must do their best to check that questionnaires were completed properly at the school. If some of them are incomplete and/or some have missing data, it may be impossible or very difficult to correct questionnaires after the initial data collection has been completed.

This is even more important for SACMEQ III than any other previous SACMEQ studies because a lot more data are being gathered – from Pupils, Teachers, School Heads, and Pupils’ Parents (or Care-takers). We are also collecting the test data on achievement in Reading, Mathematics, and Health Knowledge from pupils and teachers.

This will require the Data Collectors to check that: (a) all questions in the instruments have actually been answered, AND ALSO (b) the answers are CONSISTENT between different questions.

The Data Collectors MUST take the time to do these checks. The more effort that they put in BEFORE leaving the school, the fewer headaches NRCs will have in the data cleaning phase. Remember that the Data Collectors’ observations are the best tools for data verification.

Also the Data Collectors MUST convince the School Heads and Teachers that the study is an important one for your country and that they should therefore take the task seriously and complete the questionnaires and tests CAREFULLY.

The Data Collectors should sit with the Teachers and the School Heads if there are any problems and make sure that all tasks and checks have been completed carefully and well - BEFORE the Data Collectors leave the school.

Since there are two days of data collection, the Data Collectors should check the Teacher Booklet, School Head Booklet, and School Information Booklet during the first evening and have them fill in the missing and/or contradictory questions on the next day. For the Pupil Booklets, the Data Collectors should organize themselves so that the pupils do not leave the classroom until the Data Collectors have checked everything in the questionnaire (the second day).

CHECKS WITHIN AND ACROSS QUESTIONNAIRES

Pupil Booklet

1. Make sure that the answers given by the pupils on the following questions are consistent with the data filled in the Pupil Name Form: (a) **Pupil class** (page 5), (b) Date of birth (**PB#1**), and (c) Gender (**PB#2**). Also check that the ID boxes on **page 5, 41, 53 and 79** of the Pupil Booklet are correctly filled in.

Within Pupil Booklet checks: (Parts D and E, p.79-109)

2. If the pupil responds “No” or “Do not know” for both **PB#5** and **PB#6** (mother/father alive) then “No” should be ticked for **PB#8.1** (stays with biological parents).
3. If the pupil ticks that he / she lives with brother(s) and / or sister (s) (“Yes” is ticked for **PB#8.4**), then at least one brother (**PB#7.1**) OR one sister (**PB#7.2**) must be given.
4. If the pupil responds he / she lives in a home with family / relatives (option 1 of **PB#9**) then “yes” must be ticked for at least one of the items between **PB#8.1** and **PB#8.5**.

If the pupil responds he / she lives in a home with other people who are not members of my family (option 2 of **PB#9**) then “no” must be ticked for all the items between **PB#8.1** and **PB#8.5**.

If the pupil responds he / she lives in an orphanage or children’s home (option 4 of **PB#9**) then “no” must be ticked for all the items between **PB#8.1** and **PB#8.5** and “Yes” for **PB#8.7**.

5. Make sure that the answer to the electricity item (**PB#14.14**) is consistent with those appliances requiring electricity such as refrigerator/freezer, air conditioner, electric fan, washing machine, vacuum cleaner, computer, internet, etc. in **PB#14** as well as the electric lighting (option 5 of **PB#17**). If the majority of these items requiring electricity have ticks in “Yes”, then the pupil must have electricity at home (“Yes” for **PB#14.14**). If “No” is ticked for **PB#14.14**, then ask the pupil what the ‘power’ situation is in his or her home.
6. If the pupil has ticked “Yes” for Washing machine (**PB#14.18**) and “No” for Piped Water (**PB#14.04**), check with the pupil if he/she really has a Washing machine at home. If this is the case, then “Yes” should be ticked for Piped water (**PB#14.04**).
7. If the pupil has ticked “Yes” for Internet (**PB#14.21**) and “No” for Computer (**PB#14.20**), check with the pupil if he/she really has Internet. Ask if Internet is through a Computer (**PB#14.20**) or through Mobile (Cell) Phone (**PB#14.31**), and make sure that “Yes” is ticked for the correct one.
8. If the pupil has ticked “No” for Computer (**PB#14.20**), then “Never” should be ticked for using a computer at home (**PB#31.2**).

9. If the pupil responds that he / she does not have a mother or female guardian (option 12 of **PB#15**) and does not have not have a father or male guardian (option 12 of **PB#16**), then “No” should be ticked for **PB#8.1** and **PB#8.2**.
10. For the lighting question (**PB#17**) and the questions on the construction materials used in the living place (**PB#18**, **PB#19**, and **PB#20**), there should be only one answer. If two or more answers are given, then you should ask the pupils which is the correct answer. Emphasize that question **PB#17** asks for the main source of lighting, and **PB#18-PB#20** refer to construction materials mostly made of.
11. If the pupil has never repeated (first option in **PB#23**), this means that he/she is not repeating Grade 6 (**PB#24**). And **PB#24.1** through **PB#24.4** should be skipped. If he/she is repeating Grade 6 now, this means he/she has repeated at least once (option 2, 3, or 4 in **PB#23**). And there must be at least one “Yes” for the reasons in **PB#24.1** through **PB#24.4**. Do they agree? Ask the pupils if the answers do not agree.
12. If the pupil ticked “No” for **PB#27**, make sure that **PB#27.1** to **27.3** are blank.
13. If the pupil ticked “No” for **PB#31**, make sure that **PB#31.1** to **31.6** are blank.
14. If the pupils receive meals from school, (**PB#35**), this means that the answer to the frequency of lunch (**PB#34.2**) should be either “3 or 4 days per week” or “every day of the week”.
15. If the pupil takes extra lessons (“Yes” in **PB#36**), then there should be at least one “Yes” ticked in each of the following questions: **PB#37**, **PB#38**, **PB#39**, **PB#40**, **PB#43**, **PB#47**.
16. If the pupil claims that he/she does not get any homework (**PB#49**), then he/she should answer likewise for the question about the help that he/she gets on homework (**PB#50**), teacher correcting (**PB#51**), and teacher explaining (**PB#52**). The answer to the

homework question is likely to be consistent within class. Walk around and see if there are discrepancies.

17. Check that the pupil ticked “Yes” in **PB#57** for the item indicated in **PB#58** (best source of HIV and AIDS information).

18. If the pupil ticked “No” in **PB#63**, check that the pupil left **PB#64** and **PB#65** blank.

Compare Pupil Booklet with Pupil Name Form:

19. If the biological mother of the pupil is alive then (a) “N” (**col.17**) must be filled in **PNF**, and (b) “Yes” of **PB#5** must be ticked. If the biological mother of the pupil died then (a) “Y” (**col.17**) must be filled in **PNF**, and (b) “No” of **PB#5** must be ticked. If the pupil doesn’t know whether the biological mother is alive then (a) “U” (**col.17**) must be filled in **PNF**, and (b) “I do not know” of **PB#5** must be ticked.

20. If the biological father of the pupil is alive then (a) “N” (**col.16**) must be filled in **PNF**, and (b) “Yes” of **PB#6** must be ticked. If the biological father of the pupil died then (a) “Y” (**col.16**) must be filled in **PNF**, and (b) “No” of **PB#6** must be ticked. If the pupil doesn’t know whether the biological father is alive then (a) “U” (**col.18**) must be filled in **PNF**, and (b) “I do not know” of **PB#6** must be ticked.

21. If the pupil responds that both parents died, follow the instructions above and check that the pupil ticked “No” for staying with the biological parents (tick “No” for **PB#8.1**).

22. If the pupil responded that he/she was not absent in the previous month (**PB#21**), make sure that this information is consistent with Pupil Name Form. Also make sure that all of the reasons in **PB#22** have “No” as answers.

If the pupil was absent, at least one of the reasons in **PB#22** should be ticked “Yes”.

Compare Pupil Booklet with Teacher Booklet:

23. What do pupils say about the classroom library (**PB#25**)? This information should be consistent with the response from the Teacher Booklet (**TB#10**, **TB#11**, and **TB#12.7**).
24. Look around to see where pupils sit and write. Does it make sense with what they respond about the place of sitting and writing (**PB#29** and **PB#30**)? Does it make sense in regard to what Teachers say about the availability of the furniture in their classrooms (**TB#13.1** and **TB#13.2**)?

Compare Pupil Booklet with School Head Booklet & School Information Booklet:

25. What do pupils say about the school library (**PB#26**)? This information should be consistent for all the pupils and also with the response given in the School Head Booklet (**SHB#19.01**, **SHB#20**) and the School Information Booklet (**SIB#20**).
26. Has the pupil used a computer at school (**PB#31.1**)? If this is so, then there must be at least one computer available to pupils in the school (**SIB#19.2**).
27. Do the pupils receive free meals supplied by the school? ("Yes" ticked for **PB#35**)? This should be consistent for all the pupils and also with the response given in the School Information Booklet (**SIB#15**).
28. Has the pupil attended classes /lessons on HIV and AIDS (**PB#63**)? If this is so, then these lessons must be indicated in the School Head Booklet (**SHB#37**).

Teacher Booklet

Within Teacher Booklet checks: (Part D, p. 81-98)

1. How many years has he/she been teaching (**TB#6**)? Does this make sense in respect to his/her age (**TB#2**).
2. What are the responses to the questions on in-service training? If a Teacher wrote 0 for the number of courses (**TB#7**) and days spent (**TB#8**) for in-service training, then he/she should also select “I did not attend any in-service course” for **TB#9**.
3. Is there a class library/classroom book corner (**TB#10**)? If the answer was “Yes”, then **TB#11** must have a valid number of books, and **TB#12.7** must be “Yes”.
4. What kind of classroom furniture and teaching materials do you see around? Does your impression agree with what the Teacher claims regarding the number of sitting and writing places (**TB#13**) and the items available in the classroom (**TB#14**)?
5. If the teacher ticked “Yes” for Internet (**TB#14.17**), check that “Yes” is also ticked for Computer (**TB#14.16**).
6. Do you know if there is an education resource centre (and if not, ask the school head) which serves this school? Check the answers on **TB#23** and **TB#24**.
7. Check that the teacher ticked “Yes” in **TB#34** for the item indicated in **TB#35** (best source of HIV and AIDS information).
8. If the teacher entered “0” days in **TB#38**, check that **TB#39** and **TB#40** are blank.
9. Check that the teacher ticked “Yes” in **TB#39** for the item indicated in **TB#40** (best HIV and AIDS activity).

Compare Teacher Booklet with School Form:

10. What is the size of the English or Maths or Health class that the Teacher is teaching (TB, page 5, Identification page)? Does it match with the information in the School Form?

Compare Teacher Booklet with School Head Booklet and School Information Booklet:

11. Check that the items the teacher has access to in the school (items for which “Yes” is ticked in TB#14) matches with the items available in the school (SHB#19). Also check that TB#14.16 (computer) matches with SIBK#19.3 – computers available to staff.
12. For the question on the length of each period (TB#16), make sure that this corresponds with the same question in the School Head Booklet (SHB#13).
13. Count the number of days the teacher was absent because of bad weather (TB#21.07), security reasons (TB#21.10) and strikes (TB#21.11). The sum of these three items should be the same, or very similar, to the number of official school days lost in the School Information Booklet (SIB#16).

School Information Booklet

1. Observe the environment in which the school is located. You should be able to see if the responses regarding the distances from school to different facilities make sense. If the school is located on an island, give the direct distance to SIB#1.1 to #1.8. Otherwise, give the exact distance required to reach each place. For example, 60km by car to west plus walk 10 km to south would become 70km although the direct distance would be shorter.
2. Look at the number of teachers given in SIB#2, SIB#3, SIB#4 and SIB#5. Do the sub-components in each question add up to their respective totals? Do all the totals agree with each other?

3. Compare the total enrolment (the sum of **SIB#9.1** and **SIB#9.2**) with the number of teachers claimed above, for example in **SIB#2**. Do the two sets of numbers correspond to make a reasonable proportion?
4. Compare enrolment (**SIB#9**, **SIB#10** and **SIB#12**), number of classes (**SIB#11** and **SIB#12**) with the information on the School Form.
5. Take a look at the responses on the school inspection. If the last school inspection (**SIB#13**) was before 2006 (options 1 to 5), then the answer to **SIB#14** must be 0. If the last inspection was in 2006 or 2007 (options 6 or 7), then the answer to **SIB#14** must be higher than zero.
6. Have a close look at the number of classrooms (**SIB#17**). Do the answers make sense? You can do a rough check by “stepping out” the dimensions of the school and then making your own approximate calculation.
7. How many toilets are there (**SIB#18**) in the school? The answers to this question should agree with what was observed by one of the Data Collectors.

School Head Booklet

1. How many years has he/she been the School Head (**SHB#9**)? How many years has he/she been teaching (**SHB#8**)? Check the answers against his/her age (**SHB#5**). Do the answers make sense? If the School Head responded using a half year, it should be rounded upwards. For example 2½ years becomes 3 years.
2. How did the School Head respond to the questions on specialized training (**SHB#10** and **#11**)? If the School Head ticked “No” (option 1) in **SHB#10** or **SHB#11**, then they

should have left the box for days blank. If they ticked “Yes” (option 2) in **SHB#10** or **SHB#11**, then make sure that the box for days is filled in.

3. If the School Head does not teach, then both **SHB#12** and **SHB#13** should be 00. If they teach, but left **SHB#13** blank or put 0, then obtain the information of the length of a period from the Teacher Booklet (**TB#16**) and fill them in.
4. How did the School Head answer the ranking of activities (**SHB#14**)? Does the ranking follow the numbers 1 through 6?
5. The “wealth” and the condition of the school can be seen by looking around. Does your impression agree with the list of school resources (**SHB#19**) and the question about school buildings (**SHB#18**)?
6. If the School Head claims that he/she has a school library (**SHB#19.01**), then check that this is consistent with **SHB#20** on the borrowing policy of the school library and the number of books in the school library in the School Information Booklet (**SIB#20**).
7. Check that the School Head ticked “Yes” in **SHB#28** for the item indicated in **SHB#29** (best source of HIV and AIDS information).
8. Check that the School Head ticked “Yes” in **SHB#33** for the item indicated in **SHB#34** (best HIV and AIDS activity).

School Form and Pupil Name Form

1. Check that **Class** and **Session** (Shift) in the School Form match with **Class** and **Session** (Shift) in the Pupil Name Form. For example, if class 6B is in Session (Shift) 2 in the School Form, it should also be in Session (Shift) 2 in the Pupil Name Form.

Section 2: Special Instructions on Data Editing before Data Entry

The NRC should proceed with these operations before data entry commences.

Although some data verification procedures took place during the data collection, it would be very useful for NRCs to go through the instruments to check for internal consistency. If you notice discrepancies, then use a red pen to correct by hand directly in the instruments.

Pupil Booklet (*Parts D and E, p.79-109*)

29. The date of birth (**PB#1**) should be consistent to the answers in Pupil Name Form (PNF). If the date of birth (day and/or month and/or year) in Pupil Booklet is missing or inconsistent with PNF, copy the information from PNF (the priority is given to PNF).
30. The answers to PNF, **PB#2** (gender) and pupil sex in the ID box of Part B (p.41) are linked. If there are discrepancies, the priority is given to PNF (or to ID box of Part B if PNF missing.).

Figure 1: Examples of inconsistencies between PNF, PB#2 and ID box of Part B:

	PNF	PB#2	Id Box		PNF	PB#2	Id Box
If	B	G	B	then change to	B	B	B
If	B	G	missing	then change to	B	B	B
If	G	missing	B	then change to	G	G	G
If	missing	B	G	then change to	G	G	G
If	missing	B	missing	then change to	B	B	B

Note: B = Boy; G=Girl

31. The answers to **PB#5** (biological mother), **PB#6** (biological father), PNF, **PB#8.1** (stay with my biological parents), **PB#8.2** (stay with my guardians), **PB#15** (mother level of education) and **PB#16** (father level of education) are linked:
- If there are discrepancies between **PB#5** and PNF then the priority is given to PNF.
 - If there are discrepancies between **PB#6** and PNF then the priority is given to PNF.
 - If a pupil has ticked “No” (option 1) or “I do not Know” (option 3) to both questions **PB#5** and **PB#6**, then he/she should tick “No” (option 1) to **PB#8.1**.
 - If a pupil ticked “Yes” (option 2) for **PB#5** and **option 12 of PB#15** then he/she should tick **option 11 of PB#15** (“I don’t know”).
 - If a pupil ticked “Yes” for **PB#6** and **option 12 of PB#16** then he/she should tick **option 11 of PB#16** (“I don’t know”).
 - If a pupil ticked “Yes” (option 2) for at least one of the both items **PB#8.1** and **PB#8.2** and option 12 for both items **PB#15** and **PB#16** then change **PB#15** and **PB#16** by option 11.

Figure 2: Inconsistencies between PB#5, PB#6, PB#8.1, PB#8.2, PB#15 and PB#16:

	PB#5	PB#6	PB#8.1	PB#8.2	PB#15	PB#16		PB#5	PB#6	PB#8.1	PB#8.2	PB#15	PB#16
If	1	1					then change to	1	1	1			
If	1	3					then change to	1	3	1			
If	3	1					then change to	3	1	1			
If	3	3					then change to	3	3	1			
If	2				12		then change to	2				11	
If		2				12	then change to		2				11
If			1	2	12	12	then change to			1	2	11	11
If			2	1	12	12	then change to			2	1	11	11
			2	2	12	12				2	2	11	11

32. The answers to **PB#7.1** (# brothers), **PB#7.2** (# sisters), **PB#8.4** (stays with my brothers and/or sisters) and **PB#22.4** (take care of brothers and/or sisters) are linked. The priority is given to both questions **PB#7.1** and **PB#7.2**:
- If **PB#7.1** and **PB#7.2** are blank, then filled zeros.
 - If a pupil ticked “Yes” (option 2) to **PB#8.4** (with my brothers and/or sisters) and filled 0 for both items **PB#7.1** and **PB#7.2**, then change **PB#8.4** by “No” (option 1).

- If a pupil ticked “Yes” (option 2) to **PB#22.4** and filled 0 for both questions **PB#7.1** and **PB#7.2**, then change **PB#22.4** by “No” (option 1).

Figure 3: Inconsistencies between PB#7.1, PB#7.2, PB#8.4 and PB#22.4:

	PB#7.1	PB#7.2,	PB#8.4	PB#22.4		PB#7.1	PB#7.2,	PB#8.4	PB#22.4
If	blank				then change to	0			
If		blank			then change to		0		
If	0	0	2		then change to	0	0	1	
If	0	0		2	then change to	0	0		1

33. The answers to **PB#8** (with whom the pupil stays) and **PB#9** (where the pupil stays) are linked:

- If a pupil ticked “No” (option 1) to all the questions between **PB#8.1** and **PB#8.9**, then change **PB#8.9** by “Yes” (option2).
- If a pupil ticked **option 1 of PB#9**, then “Yes” (option 2) must be ticked for at least one of the items between **PB#8.1** and **PB#8.5**.
- If a pupil ticked **option 2 of PB#9**, then “Yes” (option 2) must be ticked for at least one of the three items **PB#8.6**, **PB#8.7** and **PB#8.8** and “No” (option 1) for the other items.
- If a pupil ticked **option 3 or option 4 of PB#9**, then “Yes” (option 2) must be ticked for at least one of the five items **PB#8.4**, **PB#8.5**, **B#8.7**, **PB#8.8** and **PB#8.9** and “No” (option 1) for the other items..

Figure 4: Inconsistencies between PB#8.1 to PB#8.9 and PB#9:

Pupil Booklet Question										
8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9	
If	1	1	1	1	1	1	1	1		then change to
If									1	then change to
If									2	then change to
If									3	then change to
If									4	then change to

Pupil Booklet Question										
8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9	
1	1	1	1	1	1	1	1	2	1	
At least one "Yes"									1	
1	1	1	1	1	At least one "Yes"			1	2	
1	1	1	At least one "Yes"		1	Or at least one "Yes"			3	
1	1	1	At least one "Yes"		1	Or at least one "Yes"			4	

34. If **PB#12** (books at home) is blank then write a zero.
35. The answers to **PB#14.14** (electricity at home), **PB#14.15** to **PB#14.31** and **option 5 of PB#17** are linked.
- If a pupil did not tick “Yes” (option2) for electricity (**PB#14.14**), then he/she should not have ticked “Yes” for items requiring electricity (**PB#14.15** to **PB#14.31**) such as refrigerator/freezer, air conditioner, electric fan, washing machine, vacuum cleaner, computer, internet, etc. Also he/she should not have ticked electric lighting (**option 5 of PB#17**). If the majority of these items requiring electricity were “Yes”, tick “Yes” for electricity (**PB#14.14**).
36. If a pupil ticked “No” (option 1) for the three items **PB#14.20** (computer at home), **PB#14.23** (TV at home) and **PB#14.31** (telephone / mobile (cell) phone), then he / she should tick “No” (option 1) for **PB#14.21** (Internet at home).
37. For the lighting question (**PB#17**) and the questions on the construction materials used in the living place (**PB#18**, **PB#19**, and **PB#20**) there should be only one answer. If two or more answers are given, cross out the lower one(s), i.e., keep the highest one. For example, if a pupil ticked “gas lamp” and “electric light” in **PB#17**, then cross out “gas lamp”. Likewise, if a pupil ticked “candle” and “there is no lighting” then cross out “there is no lighting”.
38. The answers to **PB#21** (# schools days absent), PNF and **PB#22** (reasons of the absence) are linked:
- If there is a discrepancy between **PB#21** and PNF, the priority is given to PNF.
 - If the response was given using a half day, a half day should be rounded upwards. For example, 2½ days becomes 3 days. If it’s blank, write 0.
 - If a pupil responded that he/she was not absent in the previous month (0 in **PB#21**), he/ she should skip **PB#22**.

- If a pupil filled a non-zero number in **PB#21** and did not tick “Yes” (option 2) for any of the reasons, then tick “Yes” (option 2) for **PB#22.6** (“Other reasons”).
39. The answers to **PB#23** (# grade repetition) and **PB#24** (repeating Grade 6) are linked:
- If at least one of the reasons (items between **PB#24.1** and **PB#24.4**) was ticked “Yes” (option 2), this means that he/she is repeating Grade 6. In this case, the answer to **PB#24** should be “Yes” (option 2).
 - If a pupil has ticked **option 1 of PB#23** (“never repeated”) and **option 2 of PB#24** (“Yes- I am repeating Grade 6 this year”) this means that he/she is repeating at least once. Change **PB#23** by option 2 (“I have repeated once”).
40. If a pupil ticked “Yes” for at least one of the three places “to take books home” (**PB#27.1**, **PB#27.2** and **PB#27.3**), this means that he/she is able to take books home. In this case, the answer to **PB#27** should be option 2 (“Yes- I am able to take books from”).
41. If any item on the learning materials (**PB#28**) is blank, write a series of zeros.
42. The answers to **PB#31** (the use of computer) and the items between **PB#31.1** and **PB#31.6** (the places for using a computer) are linked:
- If a pupil ticked “Never” (option 1) for all the places between **PB#31.1** and **PB#31.6** then he/she should tick for **PB#31** “No” (option 1).
 - If option 2, 3 or 4 for at least one of the places (**PB#31.1** to **PB#31.6**) was ticked this means that a pupil has used computer. In this case, the answer to **PB#31** should be changed to option 2.
43. The answers to **PB#34.1** through **PB#34.3** (meals) and **PB#35** (meals supplied by the school) are linked.
44. The answer to **PB#49** through **PB#52** on homework should be consistent. If a pupil responded that he/she does not get any homework (**PB#49**), then he/she should answer

likewise for the questions about (i) the help that he/she gets on homework (**PB#50**); (ii) the teacher correcting the homework (**PB#51**); and (iii) the teacher working through solutions on homework (**PB#52**).

45. Make sure that the best source of HIV and AIDS information filled by each pupil (**PB#58**) is ticked as “Yes” (option 2) in **PB#57**. For example, a pupil filled “16” in **PB#57**. It means that the best source of information is **PB#57.16** (Friends). Then **PB#57.16** should be ticked as “Yes” (option 2).

46. The answers to **PB#63** (attend classes/ lessons on HIV and AIDS), **PB#64**, **PB#65** and **PB#57.13** are linked. The priority is given to **PB#65** (or **PB#64** if **PB#65** missing):

- Make sure that the best HIV and AIDS activity filled by a pupil (**PB#65**) is ticked as “Yes” (option 2) in **PB#64**.
- If at least one of the activities between **PB#64.1** and **PB#64.9** was ticked as “Yes” (option 2), this means that he/she has attended classes / lessons. In this case, the answers to **PB#63** and **PB#57.13** should be “Yes” (option 2).
- If a pupil has attended classes / lessons on HIV and AIDS during this school year (**option2 of PB#63**) then he / she should tick “Yes” (option 2) to **PB#57.13** (receive HIV and AIDS information from classroom lessons).

47. If a pupil ticked “Yes” (option 2) to **PB#64.6** (a person living with HIV gave a talk during the classes / lessons) then “Yes” (option 2) should be ticked to **PB#57.22** (a person living with HIV).

48. If a pupil ticked “Yes” (option 2) to **PB#64.8** (trip to a hospital / care centre during the classes / lessons) then “Yes”(option 2) should be ticked to **PB#57.14** (Hospital / Clinic).

Teacher Booklet (*Part D, p. 81-98*)

1. The class identification and the size of the English or Maths or Health class that the Teacher is teaching (**TB, page 5, Identification page**) should match with the information in the School Form. If the class information in the Teacher Booklet is missing or different from the School Form, use the School Form information in order to correct the information in the Teacher Booklet.
2. The information about the sex of the teacher appears in two places: **TB#1** and the Identification Box for Part C: Health Knowledge Test (**page 47**). They should agree to each other. If there is a discrepancy, then use the information **TB#1** in order to correct the information in the Identification Box on **page 47**.
3. The number of years that a teacher has been teaching (**TB#6**) should make sense in relation to their age (**TB#2**). For example, it would not make sense if a teacher is 30 years old, and he/she has been teaching for 20 years. Make sure that the difference between **TB#2** and **TB#6** is around 20 (estimated age to start teaching). If this is not the case, contact the school to verify the information. Note that the number used for these questions should be whole numbers. For example, if a teacher has written 10 ½ for **TB#6**, then round it up to 11.
4. The number of courses (**TB#7**), days spent for in-service training (**TB#8**), and the effectiveness of the in-service courses (**TB#9**) are linked together. As shown in Figure 5, if consistency occurs between **TB#7** and **TB#8**, then **TB#9** should decide. If **TB#7** and **TB#8** are consistent, but disagree with **TB#9**, then you need to correct **TB#9**. If the answers to both **TB#7** and **TB#8** are greater than or equal to 1, and the answer to **TB#9** is “I did not attend any in-service course”, this option is wrong. Cross it out and leave it as ‘omitted’. This will have to be treated separately during the imputation stage. Please make sure to use the whole numbers for **TB#7** and **TB#8**.

Figure 5: Inconsistencies among TB#7, TB#8 and TB#9

	TB#7	TB#8	TB#9		TB#7	TB#8	TB#9
If	0	Non-zero	Option 1	then change to	0	0	Option 1
If	Non-zero	0	Option 1	then change to	0	0	Option 1
If	0	0	Option 2 to 5	then change to	0	0	Option 1
If	Non-zero	Non-zero	Option 1	then change to	Non-zero	Non-zero	omitted

5. The number of classroom library books (**TB#11**), the borrowing policy on the classroom library books (**TB#10**), and the availability of classroom library, book corner or book box (**TB#12.7**) are all related. If there are discrepancies, then **TB#11** should be used in order to correct the responses for **TB#10** and **TB#12.7**. Figure 6 shows the rules for corrections. If **TB#11** is 0, then **TB#10** will be automatically option 1, and **TB#12.7** will be automatically “No”. If **TB#11** is non-zero, then **TB#12.7** will be automatically “Yes”, but **TB#10** should be either option 2 or 3. If **TB#11** is non-zero and **TB#10** is option 1, then change **TB#10** to option 2. Check the answer for **TB#12.7** against the answers by the pupils who are taught by this teacher (**PB#25**).

Figure 6: Inconsistency among TB#11, TB#10 and TB#12.7

	TB#11	TB#10	TB#12.7		TB#11	TB#10	TB#12.7
If	0	whatever	whatever	then change to	0	Option 1	No
If	Non-zero	Option 1	whatever	then change to	Non-zero	Option 2	Yes

6. Access to computers (**TB#14.16**), access to Internet (**TB#14.17**), and the information about the availability of computers for staff in the School Information Booklet (**SIB#19.3**) are linked together. If computers for staff (**SIB#19.3**) are not available, then the answers to the access to computers (**TB#14.16**) and internet (**TB#14.17**) should become “No” automatically. If **SIB#19.3** is available and the internet (**TB#14.17**) is accessible, then the access to computers (**TB#14.16**) must be “Yes” automatically.

Figure 7: Inconsistencies among TB#14.16, TB#14.17, and SIB#19.3

	TB#14.16	TB#14.17	SIB#19.3		TB#14.16	TB#14.17	SIB#19.3
If	whatever	whatever	no	then change to	no	no	no
If	no	no	no	then change to	yes	no	no

7. Follow the previous logic in order to solve the problems between radio (SHB#19.21 vs TB#14.10), TV (SHB#19.23 vs. TB#14.11), and photocopier (SHB#19.29 vs. TB#14.18).
8. The number of periods to teach (TB#15) and the length per period (TB#16) are linked, and they must be written using whole numbers. If TB#16 is missing, then obtain information from SHB#12. If the School Head is not teaching, look at the instruments of the other teachers at the same school. If there are surprising results, then contact the school (and the teachers) through the data collectors.
9. If the answer to TB#23 about education resource centre is No, then the answer to TB#24 must be the first option “there is no education resource centre”. If option 2 or 3 is selected for TB#24, then change the answer for TB#23 to 2 to “Yes”. If option 1 or 2 in TB#24 is selected, then TB#24.1 through #24.6 must be blank.

Figure 8: Inconsistency among TB#23, TB#24 and TB#24.1-7

	TB#23	TB#24	TB#24.1-6		TB#23	TB#24	TB#24.1-6
If	No	Option 1 or 2	At least 1 tick	then change to	Yes	Option 3	At least 1 tick
If	No	Option 2	No tick	then change to	Yes	Option 2	No tick
If	No	Option 3	No tick	then change to	No	Option 1	No tick
If	Yes	Option 1 or 2	At least 1 tick	then change to	Yes	Option 3	At least 1 tick
If	Yes	Option 3	No tick	then change to	Yes	Option 3	omitted

10. If there are multiple ticks in TB#33, choose the highest option.

11. The teacher must have ticked “Yes” in **TB#34** for the item indicated in **TB#35** (best source of HIV and AIDS information). If this item has “No” as an answer, then change it to “Yes”. The answer to **TB#35** will be followed.
12. The number of days of HIV and AIDS specialized training (**TB#38**), activities during these lessons (**TB#39**), the best activity (**TB#40**), and information on HIV and AIDS received during in-service teacher training (**TB#34.14**) are all linked. If there are inconsistencies between **TB#38** and **TB#34.14**, use the information in **TB#38** to correct the answer to **TB#34.14**. If the item used in **TB#39** did not have “Yes” as an answer in **TB#38**, then change it to “Yes”.
13. Some of the sub-items in **TB#34** and **TB#38** are related. For example, if a person living with HIV gave a talk during the in-service course (**TB#38.07**), then the information received from a person living with HIV (**TB#34.23**) must be automatically “Yes”, but not the other way around because the information may have come from outside the in-service course. Another example is “we had an organized trip to a hospital / care centre” (**TB#38.09**) and information received from hospital / clinic (**TB#34.15**). Look at also between “we watched a video / film” (**TB#38.04**) and information received from video player (VCR, DVD, etc.) (**TB#34.03**). The rule is that if the item is ticked “Yes” in **TB#38**, then it must also be ticked “Yes” in **TB#34**.

School Head Booklet

1. Compare the level of education provided by the school (**SHB#2**) with the nearest school offering secondary education as indicated in the School Information Booklet (**SIB#1.5**). If **SIB#1.5** is "0", then for **SHB#2** either option 3 or 4 should be ticked. If this is not the case, you should contact the school to get the correct information.
2. For the number of years of teaching (**SHB#8**), if a School Head responded using a half year, it should be rounded upwards. For example, 10½ years becomes 11 years. Likewise, a half year in **SHB#9** should be rounded upwards. Number of days in **SHB#10** (option 2) and **SHB#11** (option 2) should also be rounded upwards.
3. The number of years being a School Head (**SHB#9**) should be less than or equal to the number of years teaching (**SHB#8**). If not, contact the school head and get the correct information.
4. The number of years that the School Head has been teaching (**SHB#8**) should make sense in relation to their age (**SHB#5**). For example, it would not make sense if a School Head is 30 years old, and he/she has been teaching for 20 years. Make sure that the difference between **SHB#5** and **SHB#8** is around 20 (estimated age to start teaching). If this is not the case, contact the School Head to verify the information.
5. If the School Head ticked "No" (option 1) of **SHB#10** but also filled in the boxes for how many days of special training received, then cross out "No" and tick "Yes" (option 2). If the second option is ticked but the boxes for how many days are blank or filled with 00, then cross out "Yes" and tick "No" (option 1). Do the same for **SHB#11**.

6. If the School Head does not teach (zero or blank in **SHB#12**), it could be possible that he/she left **SHB#13** blank. If **SHB#13** is missing or zero, then obtain the information of the length of a period from the Teacher Booklet (**TB#16**) and fill them in.

Figure 9: Inconsistency between SHB#12 and SIB#13

SHB#12	SHB#13		SHB#12	SHB#13
If Non-zero	0	then change to	Non-zero	Copy from TB#16

49. The answers to **SHB#19.20** (electricity), and **SHB#19.21** to **SHB#19.32** are linked. If the School Head did not tick “Yes” for electricity (**SHB#19.20**), then he/she should not have ticked “Yes” for items requiring electricity (**SHB#19.21** to **SHB#19.32**) such as TV, fax machine, photocopier, computer(s), etc. If the majority of these items requiring electricity were “Yes”, tick “Yes” for electricity (**SHB#19.20**).

7. Check the agreement between the number of books in the school library in the School Information Booklet (**SIB#20**), the availability of a library (**SHB#19.01**), and the policy on borrowing library books (**SHB#20**). If there are no library books (“0” in **SIB#20**), then make sure that “No” is ticked for library (**SHB#19.01**) and “We have no school library” (option 1) is selected for **SHB#20**. If there are library books (non-zero value in **SIB#20**) then make sure that “Yes” is ticked for library (**SHB#19.01**), and if “We have no school library” (option 1) is selected for **SHB#20**, change this one to “No” (option 2). An exceptional case is when “No” is ticked for library in **SHB#19.01** and “We have no school library” (option 1) is ticked for **SHB#20**, but **SIB#20** has a non-zero value. In this case, change **SIB#20** to “0”.

Figure 10: Inconsistency among SIB#20, SHB#19.01 and SHB#20

	SIB#20	SHB#19.01	SHB#20		SIB#20	SHB#19.01	SHB#20
If	0	No	Option 2 or 3	then change to	0	No	Option 1
If	0	Yes	Option 1	then change to	0	No	Option 1
If	0	Yes	Option 2 or 3	then change to	0	No	Option 1
If	Non-zero	No	Option 1	then change to	0	No	Option 1
If	Non-zero	Yes	Option 1	then change to	Non-zero	Yes	Option 2
If	Non-zero	No	Option 2 or 3	then change to	Non-zero	Yes	Option 2 or 3

8. If **SHB#29** contains a valid number (between 1 and 24), then this item must be ticked as “Yes” in **SHB#28**.

9. If **SHB#34** contains a valid item number (between 1 and 14), then this item must be ticked as “Yes” in **SHB#33**.

10. If any of the items in **SHB#33** are ticked as “Yes”, then **SHB#28.14** should be ticked “Yes” (in-service teacher training), and **SHB#11** should also be ticked “Yes” (specialized training received in health, safety, etc.). If the number of days in **SHB#11** is left blank, contact the School Head to get the correct information.

11. Some of the sub-items in **SHB#28** and **SHB#33** are related. For example, if a person living with HIV gave a talk during the in-service course (**SHB#33.07**), then the information received from a person living with HIV (**SHB#28.23**) must be automatically “Yes”, but not the other way around because the information may have come from outside the in-service course. Another example is “we had an organized trip to a hospital / care centre” (**SHB#33.09**) and information received from hospital / clinic (**SHB#28.15**). Look also between “we watched a video / film” (**SHB#33.04**) and information received from video player (VCR, DVD, etc.) (**SHB#28.03**). The rule is that if the item is ticked “Yes” in **SHB#33**, then it must also be ticked “Yes” in **SHB#28**.

12. If any of the sub-components for the number of pupils and teachers (**SHB#35**) is blank, write 000.

Check whether **SHB#38.01** to **SHB#38.15**, regarding kind of support given to pupils/staff, are consistent. If any of the boxes under “With the assistance of...” is ticked, then the box “Yes” must be ticked under “This happened at my school this year”. If none of the boxes under “With the assistance of...” is ticked, then the box “No” must be ticked under “This happened at my school this year”.

School Information Booklet

1. If any of the sub-components for the number of teachers (**SIB#2, #3, #4, #5, #6, #7, #8**) is blank, fill the boxes with “0”s. The sub-components should add up to the total number of teachers in each question. Contact the school to get the correct information if this is not the case.
2. The total number of male teachers, female teachers, and total number of teachers (male and female) should also agree with each other. If any of these do not agree with each other, you should call the school and get the correct information.
The following calculations should be made:

Male teachers:	SIB#2.1 + SIB#2.2 = SIB#5.06 + SIB#5(b) = School Form (2007) i.e. (permanent + non-permanent male teachers) = (absent + non-absent male teachers) = total nr of male teachers in the School Form.
Female teachers:	SIB#2.3 + SIB#2.4 = SIB#5.12 + SIB#5(d) = School Form (2007)
Total nr of teachers:	SIB#2.5 = SIB#3.6 = SIB#4.7 = SIB#5(e) = School Form (2007)

3. The answer to **SIB#8** should be less than the answer to **SIB#6**. The answer to **SIB#8** should also be less than the answer to **SIB#7.1**. If this is not the case, contact the School Head and get the correct information.

4. The Grade 6 enrolment in **SIB#10** and Grade 6 classes in **SIB#11** should match the information on the **School Form**. If this is not the case, contact the School Head and get the correct information.
5. The sum of the number of pupils in **SIB#12** (add **SIB#12.1** to **SIB#12.5**, column 1) should be the same as the total nr of pupils in the **School Form** (for 2007), and also the total nr of pupils indicated in **SIB#9**. Likewise, the total number of **classes** (add **SIB#12.1** to **SIB#12.5**, column 2) should be the same as the total nr of classes in the **School Form**. If this is not the case, contact the School Head and get the correct information.
6. If the last school inspection (**SIB#13**) was before 2006 (options 1 to 5), then the answer to **SIB#14** must be 0. If **SIB#14** is 0 then option 6 or 7 should not be selected in **SIB#13**. If this information is incorrect, you should call the school to get the correct information.
7. If any of the boxes in the number of classrooms (**SIB#17**), number of toilets (**SIB#18**), number of computers (**SIB#19**), and number of books in the library (**SIB#20**) are blank, then write a zero.
8. If any of the sub-components for number of computers in the school (**SIB#19.1 – 19.4**) is non-zero, then **SHB#19.31** (availability of computers in the school) should be 'Yes'.

Section 3: Special Coding Instructions during Data Entry

<p><u>The NRCs</u> should proceed with the first part.</p>
--

Installation of WinDEM

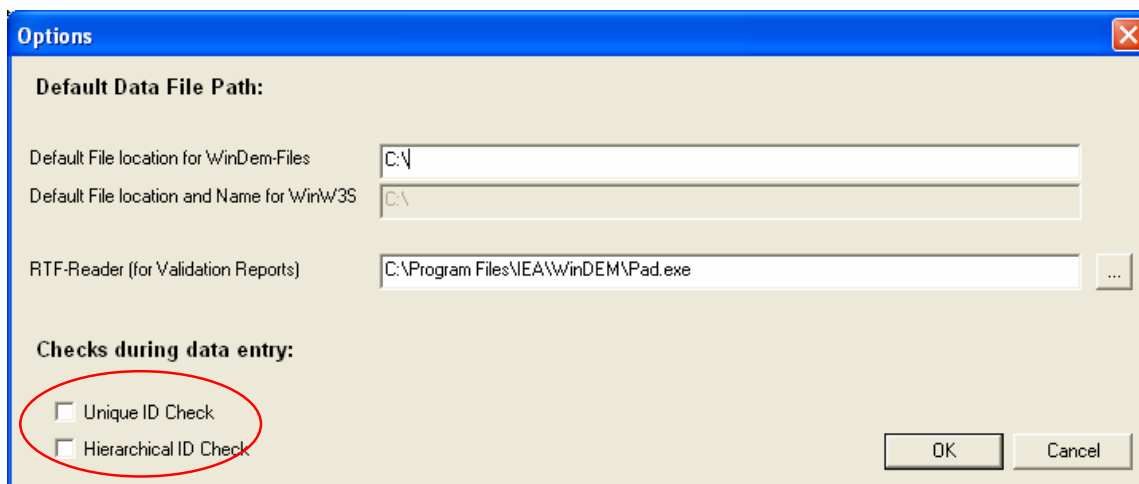
You have received an installation file for WinDEM by e-mail. Please install this software on all the computers that you will be using for the SACMEQ3 data entry. Simply double click on the file “**WinDEM_3_2_20_3.exe**”, and follow the instructions using the default settings.

In order to check that WinDEM works properly, try to enter into the WinDEM software with the following user name and the password:

User Name: admin
Password: cookie

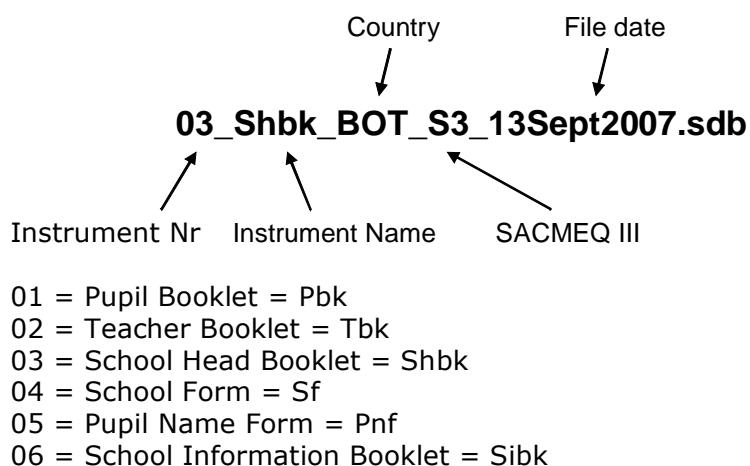
Make sure that only the upper case is accepted here. The above user name and the password are for you to use, not for the data punchers. The first thing that you will need to do is to change certain items in the options. Select the Options menu, and come to the “Setting...” submenu. At the bottom of this setting window, please make sure that **two boxes are not ticked**. If they are ticked, de-click them so that these boxes would be empty as shown in Figure 11.

Figure 11: Options to be changed in WinDEM



WinDEM Codebooks

The following is an explanation of how the WinDEM codebook files are named, using Botswana as an example.



For the Botswana example, the files are named as follows:

Pupil Booklet:

- 01_PbK_PartA_(Read)_BOT_S3_13Sept2007.sdb
- 01_PbK_PartB_(Hlth)_BOT_S3_13Sept2007.sdb
- 01_PbK_PartC_(Math)_BOT_S3_13Sept2007.sdb
- 01_PbK_PartD_(Qnre)_BOT_S3_21Sept2007.sdb - (includes Part E)

Teacher Booklet:

- 02_TbK_PartA_(Read)_BOT_S3_11Sept2007.sdb
- 02_TbK_PartB_(Hlth)_BOT_S3_13Sept2007.sdb
- 02_TbK_PartC_(Math)_BOT_S3_11Sept2007.sdb

- 02_Tbk_PartD_(Qnre)_BOT_S3_13Sept2007.sdb

School Head Booklet:

- 03_Shbk_BOT_S3_13Sept2007.sdb

School Form:

- 04_Sf_BOT_S3_13Sept2007.sdb

Pupil Name Form:

- 05_Pnf_BOT_S3_13Sept2007.sdb

School Information Booklet:

- 06_Sibk_BOT_S3_13Sept2007.sdb

Data Entry arrangement

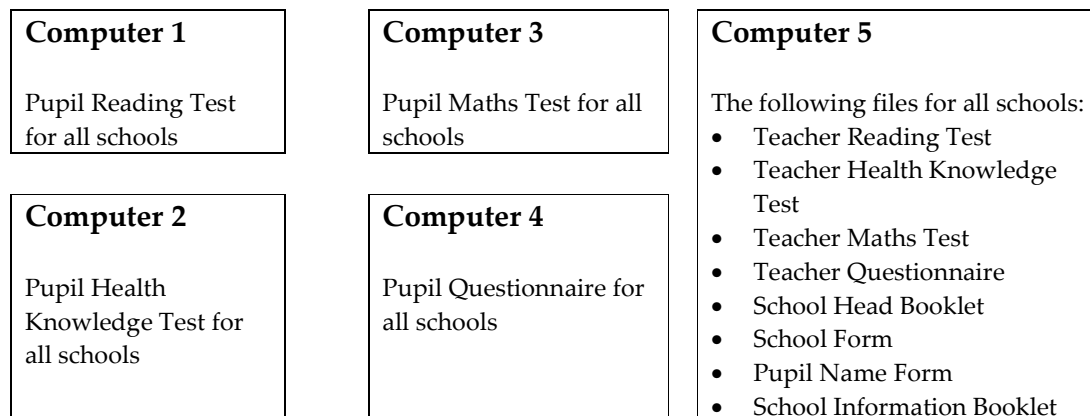
It would be very helpful at this time to make a diagram illustrating which computers will be used to enter which instruments. For example, you may use a computer for each region. In this case, you may make a diagram such as (see Figure 12):

Figure 12: Diagram of use of computers 1

Computer 1	Computer 2	Computer 3	Computer 4	Computer 5
Schools 1-15 Schools 16-30	Schools 31-45 Schools 46-60	Schools 61-75 Schools 76-90	Schools 91-105 Schools 106-120	Schools 121-135 Schools 136-150

On the other hand, you may choose to dedicate four computers for pupil booklet of all the schools and one computer for the other instruments for all the schools. In this case, your diagram would look like (see Figure 13):

Figure 13: Diagram of use of computers 2



This would reduce the confusion when compiling files from different computers (see Step 1 in Section 4).

Try entering some data for yourself

Instead of leaving all the tasks to your Data Punchers, it would be advised to start the entry of one of the schools (for all the instruments) on your own. If you put yourself into the shoes of Data Punchers, you will be more prepared to answer to questions by them during the data entry of the rest of the data.

Data Punchers should proceed with the operations below.

General Instructions

• **Different parts of the booklet**

For SACMEQ III, there are many files to deal with. Although the paper version of Pupil Booklet and Teacher Booklet have different “parts” together as a booklet, these parts have been physically separated on the electronic version. It is important to enter not only the Instrument codes but also the codes associated with the parts.

- **User profile**

For data entry, use the following User Name and Password:

User Name: staff
Password: bird

Notice that the lower case should be used for them.

- **General codes**

Enter the code which is written in the box which has been ticked. In the example below, you should type 2 (see Figure 14). Then the cursor will go to the next field automatically.

Figure 14: An example of a tick



In most cases, you do not need to hit an “ENTER” key to go to the next field. However, for numerical fields, if a response did not use the entire available number of columns, then this entry must be followed by an “ENTER” key to go to the next field. In the example below (see Figure 15), for a field with 2-digit allocation, only 1 digit is used. In this case, after typing 4, you need to press an “ENTER” key. Otherwise, the cursor will stay in this field.

Figure 15: An example of an incomplete numerical field



Instructions for Missing Responses

Since the National Research Coordinator has already verified data before data entry commences, there should be no missing information in the instruments. However, if you

do encounter some missing information, please follow the rules below unless otherwise indicated specially:

- Questions which have been administered but contain no information should be coded as 9 (if the response requires 1 digit), 99 (if the response requires 2 digits), 999 (if the response requires 3 digits), etc.
- Questions which have not been administered should be coded as 8 (if the response requires 1 digit), 98 (if the response requires 2 digits), 998 (if the response requires 3 digits), etc. If the whole instrument was missing, use this code.

Coding Rules for Booklets

(a) Pupil Booklet

1. **Multiple response code for Pupil Reading, Mathematics, and Health Tests** – Please make sure to enter the code 7 when more than one response is given as an answer for any test item. This does not apply to the questions in the questionnaires.
2. **Check digits in Pupil Reading, Mathematics, and Health Tests** – Please note that a “check digit” will need to be inserted at several stages in the tests. These check digits are to make sure that data enterers are ‘awake’. Only the specified letter should be entered for each check digit.

(b) Teacher Booklet

1. For the Teacher ID section, if any of the grade 6 English or Math or Health classes is left blank, then enter “1” (for no) for the classes, and then enter “98” for the class size.

2. **Multiple response code for Teacher Reading, Mathematics, and Health Tests** – Please make sure to enter the code 7 when more than one response is given as an answer for any test item.
3. **Check digits in Teacher Reading, Mathematics, and Health Tests** – Please note that a “check digit” will need to be inserted at several stages in the tests. These check digits are to make sure that data enterers are ‘awake’. Only the specified letter should be entered for each check digit.
4. Please pay a special attention to filter questions, such as TB#24 or PB#31. For example, if the response to TB#24 is “1” or “2”, then enter a code “8” for TB#24.1 to TB#24.6.

(c) School Information Booklet

1. For the lines not used for the session information (SIB#12), enter the not-administered codes. For example, if there is only one session which exists in school, the lines for 2nd through 6th sessions will be coded 9998 for the number of pupils per session and 98 for the number of classes per session.

(d) School Head Booklet

1. If option 1 is selected for SHB#10 and SHB #11, then the boxes at the end of option 2 will have a code “98”. If option 2 is selected, but the boxes are left blank, then it will have a code “99”.
2. Be extra careful with questions SHB#38.01 to SHB#38.15. Note that they consist of five data entry fields and not six (see Figure 16). For field A enter the response to “This happened at my school this year”. Enter ‘1’ if the response is “No” or ‘2’ if the response is “Yes”. For fields B, C, D and E the coding is different. You need to enter code ‘1’ if the box is not ticked, or enter code ‘2 if the box has been ticked.

Figure 16: Question 38 in School Head Booklet

		A		B		C		D		E	
		This happened at my school this year		With the assistance of							
		No	Yes	Ministry (Dep't) of Education	Ministry (Dep't) of Health	Other Govern't Org'ns	Other Groups ¹				
Support to pupils											
38.01	Guidance / <u>counseling</u> for orphans and vulnerable pupils	(1)	(2)	(2)	(2)	(2)	(2)	(2)			

The special codes have been summarized in Table 1.

Table 1: Special codes

Code	When to use
7	<u>Multiple response code</u> : Use this when more than one option is ticked in test items.
9, 99, 999, etc.	<u>Missing response code</u> : Use this when no response is given from the administered respondents
8, 98, 998, etc.	<u>Not-administered code</u> : Use this when the question was not encountered by the respondent.
A, B, C, etc.	<u>Check letter code</u> : Use this when check letter needs to be inserted during any test.

Section 4: Special Instructions on Data Verification after Data Entry

The NRC should proceed with these operations after the data entry has completed.

Before you proceed with data verification, it would be very helpful to go back to your original diagram illustrating which computers have been used for data entry.

Most of the procedures in this section are NOT allowed for data enterers since they need to be operated only by the NRCs only. Therefore, you will need to exit WinDEM, and re-start it with a different profile.

This time, type in the following username and password:

User Name: admin
Password: cookie

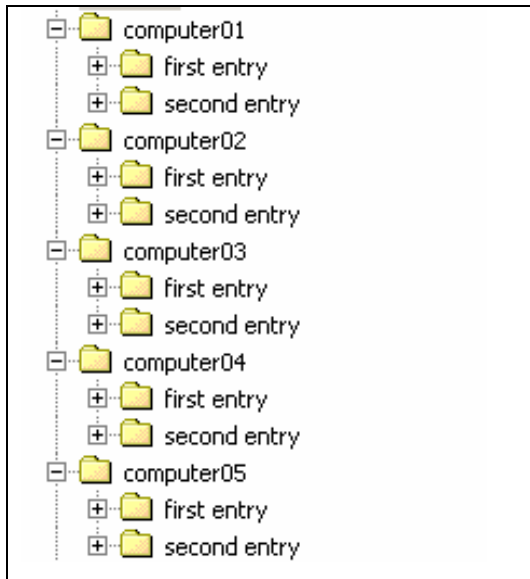
Notice again that the lower case should be used for them.

Step 1: Collect files from different computers

After entering data using several computers, you need to gather data entered on different computers to make a single file for each instrument for the first and the second entry each.

First, on a diskette or USB key, you need to make several sub-directories (representing different computers). Since you did double-entry of data, you need to make two sub-directories within each sub-directory: (1) first entry; and (2) second entry. Now you should have a structure of sub-directories as follows (see Figure 17):

Figure 17: Sub-directories to be created to copy different files



Now you need to go around all the computers that were used for the data entry. You should copy both SDB and DBF files in the corresponding sub-directories.

Step 2: Create one file for each instrument

You should now decide one computer as the parent computer, say, computer01. On this computer, you can first open a codebook. For example, you may start with the codebook for the Pupil Health Knowledge Test (01_pbk_partb_hlth_zan_s3_final.sdb). Then you can go to the Import/Export menu, and import WinDEM. You should change the sub-directory to computer02 in order to import the WinDEM data from computer02. Continue importing the WinDEM data until you finish importing from all the sub-directories for the first entry.

After you complete with this particular file (instrument), you should do the same for the other instrument. At the end, you will have twelve (12) combined datafiles. This means that you will have:

- 1 combined file for Pupil Reading Test (# of records = # of pupils surveyed),

- 1 combined file for Pupil Health Knowledge Test (# of records = # of pupils surveyed),
- 1 combined file for Pupil Mathematics Test (# of records = # of pupils surveyed),
- 1 combined file for Pupil Questionnaire (# of records = # of pupils surveyed),
- 1 combined file for Teacher Reading Test (# of records = # of teachers surveyed),
- 1 combined file for Teacher Health Knowledge Test (# of records = # of teachers surveyed),
- 1 combined file for Teacher Mathematics Test (# of records = # of teachers surveyed),
- 1 combined file for Teacher Questionnaire (# of records = # of teachers surveyed),
- 1 combined file for School Information Booklets (# of records = # of schools surveyed),
- 1 combined file for School Head Booklets (# of records = # of schools surveyed),
- 1 combined file for School Form (# of records = # of existing classes in all the surveyed schools), and
- 1 combined file for Pupil Name Form (# of records = # of pupils surveyed)

These datafiles should be used to do the rest of the cleaning procedures within WinDEM. The same procedures should be carried out for the second set of data. For each set, please follow the Steps described below.

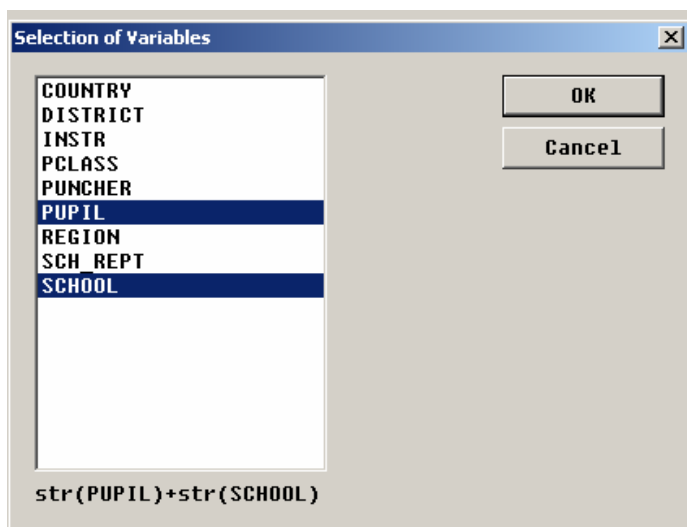
Step 3: Data Cleaning using WinDEM

✓ Unique ID Check

In all of the datafiles for the Pupil Booklet, there must be only one record for each pupil; if there is more than one record with the same pupil identification code, the program can detect this error. To run this procedure, you will choose “Unique ID” under the “Verify” menu. For all of the datafiles for the Pupil Booklet, you need to select SCHOOL (ID of school) and

PUPIL (ID of pupil) as the variables to be checked for the uniqueness of the record (see Figure 18).

Figure 18: Selection of variables in Unique ID Check for the datafile on Pupil Reading Test



For all of the datafiles for the Teacher Booklet, the unique identification would be SCHOOL (ID of school) and TEACHER (ID of teacher); and for the School Head Booklet and the School Information Booklet, it would be SCHOOL (ID of school) only. Be careful with School Form, where the unique identification would be SCHOOL (ID of school) and SFCLASS (the class code). When you are ready to run this check, click "OK". Click "Save" to save the check report in the same directory as the datafile. It is saved under the file name with a suffix "_RE1.rtf". If there is no error, the program will come back with the message "No problems found". When you click "OK", you will get another message "Do you want to open the Report right now?" You can open it immediately to see which value of the variables is repeated within a file. Close this report when you finish.

Figure 19 is an example of such a report.

Figure 19: Example of a Unique ID Check Report

Record	25	School 10	Pupil 3
Record	150	School 10	Pupil 3

The first thing to check would be the raw instruments in order to identify which record is the “true” School 10 Pupil 3. However, for the “false” record, the cleaning action depends on what type of errors. Errors could have happened because of: (i) pure duplicates; (ii) error in the Pupil ID; or (iii) error in the School ID.

Pure duplicates means that the puncher actually entered the same booklet twice. When you examine the two records on the WinDEM screen, you should be able to see that they have the same values for other variables. In this case, “false” record must be deleted.

If the error was in the Pupil ID, within the same school, one of the Pupil IDs must be missing. Identify which Pupil ID is missing in this school, and change the Pupil ID of the “false” record to the missing Pupil ID.

If the error was in the School ID, in one of the other schools in the survey, there must be a missing Pupil ID 3. Identify which School is missing Pupil ID 3, and change the School ID of the “false” record.

✓ Merge Check

This check detects records in a datafile that do not have matching records in a related datafile even for a higher level of data aggregation. For example, you can check if each pupil in the Pupil Questionnaire file has a matching school head in the School Head Booklet file. To do this, with the Pupil Questionnaire file open, first select “Merge Check” from the “Verify” menu. Then select SCHOOL (ID of school) as the linking variable. Click “OK” and then select the School Head Booklet as the file to be linked. The Report for this

check is saved with a suffix “_RE5.rtf”. If there are problems, find out in the report and identify which school(s) do not have data of School Head. Likewise, you may want to see if each pupil in Pupil Questionnaire file has a matching teacher in the Teacher Questionnaire file using SCHOOL as the linking variable.

Let us think of an example (as shown in Figure 20) of a merge check report (for merging the School Head Booklet to the Pupil Name Form):

Figure 20: Example of a Merge Check Report

Expression:	str(SCHOOL)

Record	1 '1' not found!
Record	2 '2' not found!
Record	4 '4' not found!
Record	5 '5' not found!
Record	6 '6' not found!
Record	7 '7' not found!
Record	8 '8' not found!
Record	9 '9' not found!
Record	10 '10' not found!
10 Records analysed	
9 Problem(s) found!	

The report says that for the 9 schools in the School Head Booklet, there are no matching records in Pupil Name Form. This means that in the Pupil Name Form, these schools could be actually missing or have been given a wrong School Identification Code.

✓ Double Coding Check

Since you have done the independent double data entry, you can now compare the first set of data entry with the second set. The data in the first set are supposed to be exactly the same as that in the second set. The Double Coding Check in WinDEM detects discrepancies in each of the variables. Before comparing the two sets of data entry, make sure that one of the sets has different filenames. For example, in the first set, the Pupil

Reading Test file can be named PRT1, but for the second set, it could be named PRT2. With PRT1open, select “Double Coding Check” from the “Verify” menu, then select as many variables as you wish. Select PRT2 as the file to be compared. The Report for this check is saved with a suffix “_RE6.rtf”. Since different punchers entered data on different computers, usually the errors picked up by WinDEM are the discrepancies in puncher codes. If these are the only discrepancies, then this means both files have the same data.

However, if you see the real discrepancies, then some actions are required. Think of the example (shown in Figure 21) of the two files of the Pupil Questionnaire:

Figure 21: Example of a Double Coding Check Report

Record (1) 160 pbyear = 92 Record (2) 85 pbyear = 94
--

The report indicates that Record 160 in the first file and Record 85 are supposed to be the same person (defined by School ID and Pupil ID). However, it also shows that the variable pbyear do not have the same value between the two versions of the files. Now you need to go back to the raw instrument of this given person of the given school, and check the correct figure. Then change the incorrect one. For this reason, it is crucial to carry out the Unique ID Check first.

✓ Validation Check

After correcting errors based on the Double Coding Check, the two files will be exactly the same (except for the puncher ID). For Steps 6 and 7, you can only choose one of the sets.

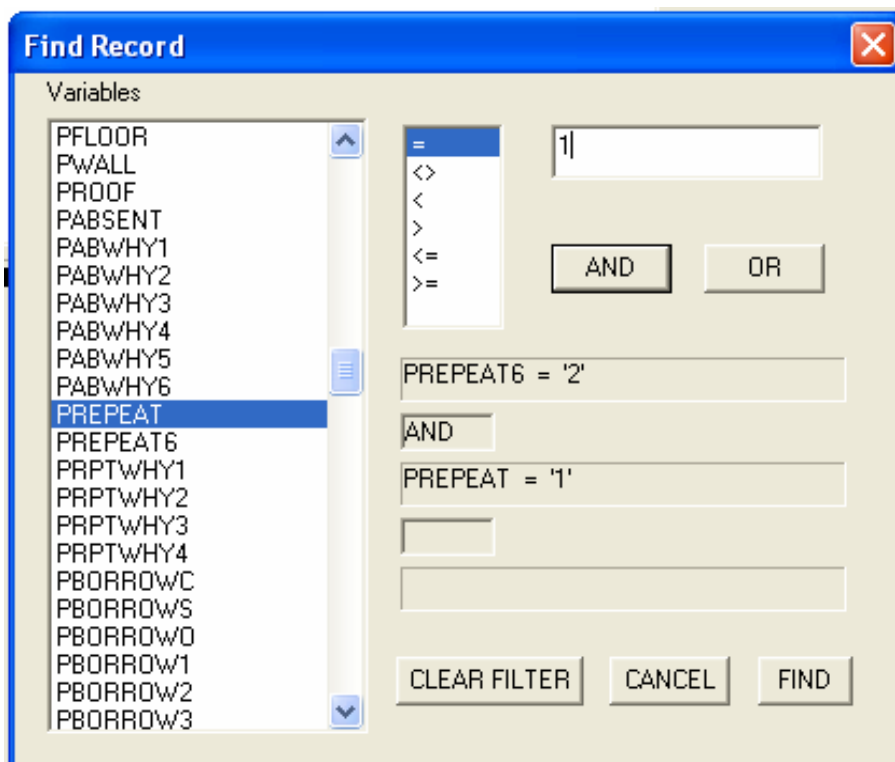
The Validation Check detects cases where data do not fall in the range validation criteria specified in the WinDEM Codebook. From the “Verify” menu, select “Validation”. In WinDEM, only those values in non-categorical variables that have been entered explicitly (by confirmation) will be detected. The Report for this check is saved with a suffix “_RE3”. You

will need to check the raw instrument but it could be possible that you would not need to correct them. For example, a School Head could be really 80 years old. This case will be flagged by the WinDEM, but you may leave it as is.

✓ Using the “Find Record(s)” operation in WinDEM

You can check the within-file consistencies using the “Find Record(s)” operation that can be found under the “Edit” menu. For example, pupils who are repeating Grade 6 (PQ#24) should not have repeated grades (PQ#23). In order to check this consistency, you can enter the **criteria for the inconsistent situation** and click “Find” as shown below (see Figure 22).

Figure 22: WinDEM Find Record operation



WinDEM would look for records that meet the criteria, and you should get the WinDEM message “No correspondence found”. If WinDEM shows some records, this means that these records contain inconsistencies between these questions. You should identify which

pupils from which schools the records are from by viewing the Table entry as shown in Figure 23, then check against the instruments.

Figure 23: Table Entry of WinDEM

	PUNCHER	COUNTRY	REGION	DISTRICT	SCHOOL	PUPIL
1	GE	TAN	DAR	ILA	3	18
2	GE	TAN	DAR	ILA	2	14
3	GE	TAN	DAR	ILA	1	20
4	GE	TAN	DAR	TMK	4	8

Please note that within WinDEM, up to three criteria could be entered connecting with “AND” or “OR”. It should be also noted that these consistency checks only apply to variables within the same file. The consistency checks between files are not available through WinDEM. They will be carried out at a later stage on a different software.

You can use this same procedure in order to check other consistencies that have been mentioned in Section 2. For example, you can test the criteria listed in Table 2. **Remember that you should be aiming for “no matches”.**

Table 2: Some criteria to be used for the Find operation of WinDEM

Meaning of Inconsistency	Syntax in WinDEM
Pupil has no electricity (PB#14.14) but has an electric fan (PB#14.17).	PPOS14=1 AND PPOS17=2 (Also try with other electric appliances.)
Pupil had no absence (PB#21), but the first reason (PB#22.1) was ticked Yes.	PABSENT=0 AND PABWHY1=2 (Also try with other reasons up to PABWHY7=2.)
Pupil receives no homework (PB#49), but he/she is helped on homework at home (PB#50).	PHMWK=1 AND PHMWKHLP<>1 (Also try with PB#51 and PB#52.)
Pupil receives homework based on PB#49 but does not receive homework based on PB#50.	PHMWK<>1 AND PHMWKHLP=1 (Also try with PB#51 and PB#52.)
Teacher had no in-service course (TB#7) but had some days for in-service course (TB#8).	TINSERV=0 AND TINSERVD<>0 (Now try with TB#9. What would be the value?)

Meaning of Inconsistency	Syntax in WinDEM
Teacher had some in-service courses (TB#7) but had 0 days for in-service course (TB#8).	TINSERV<0 AND TINSERVD=0 (Now try with TB#9. What would be the value?)
Teacher responded that there was no book in the classroom library (TB#11) but ticked 2 or 3 on the book borrowing policy (TB#10).	TBOOKCLS=0 AND TBORROW<1
Teacher responded that there was no book in the classroom library (TB#11) but ticked Yes for the classroom library resource item (TB#12.07).	TBOOKCLS=0 AND TRESCLIB=2
Teacher responded that there were some books in the classroom library (TB#11) but responded “no classroom library” in TB#10.	TBOOKCLS<0 AND TBORROW=1
Teacher responded that there were some books in the classroom library (TB#11) but responded “no classroom library” in TB#12.07.	TBOOKCLS<0 AND TRESCLIB=1
Teacher responded that there was no resource centre nearby (TB#23) but ticked 2 or 3 about visiting the resource centre (TB#24).	TRCENTRE=1 AND TRCVISIT<1
Teacher responded that there was no resource centre nearby (TB#23) but the first use was ticked as Yes (TB#24.1).	TRCENTRE=1 AND TRCUSE1=2 (Also try with other uses up to TRCUSE7=2).
Teacher responded that there was a resource centre nearby (TB#23) but responded “no resource centre” in TB#24.	TRCENTRE=2 AND TRCVISIT=1
School Head responded that the most recent inspection was 2005 or before (SIB#13), but some inspections took place since 2006 based on SIB#14.	SYINSP<6 AND SINS2006<0
School Head responded that the most recent inspection was 2006 or later (SIB#13), but no inspection took place since 2006 based on SIB#14.	SYINSP>5 AND SINS2006=0
School Head responded that he/she did not teach (SHB#12) but responded the length of a course (SHB#13).	SPERIODS=0 AND SMINUTES<0
School Head responded that he/she teaches courses (SHB#12) but put zero for the length of the course (SHB#13).	SPERIODS<0 AND SMINUTES=0
School Head ticked No for the resource item on the school library (SHB#19.01) but ticked 2 or 3 on the book borrowing policy (SHB#20).	SRES01=1 AND SBORROW<1
School Head ticked Yes for the resource item on the school library (SHB#19.01) but ticked 1 “no school library” based on SHB#20.	SRES01=2 AND SBORROW=1

Please note that the items listed in Table 2 are only partial examples. There are other inconsistencies within files to be checked, as presented in Section 2 of this document. Try to follow the syntax in Table 2, and now you can try some syntax in WinDEM.

✓ Column Check

There is another check available in WinDEM, which is called Column Check.

Column shift should not occur if you have followed the directions of entering the check digits. You can see that you have entered the appropriate check digits by looking at the “Table view” of the data (in the View menu → Table Entry → Whole Table). Make sure that there is a vertical line of “A”, “B”, “C” and so on in the Reading, Mathematics, and Health knowledge tests as shown in Figure 24.

Figure 24: Table view in WinDEM

	PREAD11	PREAD12	PREAD13	CHECKA	PREAD14	PREAD15	PREAD16	PREAD17
1	2	3	2	a	4	1	4	
2	2	1	2	a	4	1	4	
3	4	3	2	a	4	1	3	
4	4	3	2	a	4	1	3	
5	2	1	2	a	4	1	2	
6	2	1	2	a	3	1	1	
7	2	1	1	a	2	1	3	
8	2	3	2	a	1	2	1	
9	4	1	2	a	4	1	2	
10	4	1	1	a	2	1	4	
11	4	4	4	a	3	2	2	
12	1	1	2	a	1	1	1	
13	1	1	1	a	1	1	2	
14	1	1	2	a	4	1	4	
15	4	1	2	a	1	1	4	
16	2	4	1	a	3	1	4	
17	3	1	2	a	1	1	1	
18	4	1	2	a	4	1	4	
19	4	1	2	a	4	1	3	
20	2	1	2	a	1	3	3	
21	2	1	2	a	3	1	4	
22	8	8	8	a	8	8	8	
23	4	3	2	a	4	1	1	
24	4	1	2	a	4	1	1	
25	1	3	1	a	1	2	1	
26	1	3	2	a	2	1	1	

Another way to verify that there is no column shift is to use the “Column Check” which is available from the “Verify” menu. The Report for this check is saved with a suffix “_RE2.rtf”. If the correct check digits are not entered for the appropriate check variables, you should identify which pupil or teacher it is and check the Pupil Booklet or Teacher Booklet of this respondent.

Step 4: Make backup copies of data files

When you have completed the data entry and finished the verification process mentioned above to correct any errors in the data, you will need to make a copy of the data files to be sent to IIEP (either by e-mail or on diskette/CD). Use Windows Explorer or another file management software in order to make copies.

Make a copy of this diskette/CD which you can store in case the first one gets lost or reaches the IIEP in a damaged condition. If you have made backup copies of data files on diskette, lock the diskette by moving the “write protect” notch to the locked position.

Step 5: Export the codebook and the datafile into SPSS

If you are sure that your data are clean, and if you want to treat the data on SPSS, you can export your WinDEM information to SPSS while within WinDEM. This procedure is not required for the SACMEQ NRCs.

In order to generate the control file for SPSS, you can choose “Open Codebook” from the File menu. Select the codebook you wish to export. The software will beep because the data already exist. Choose “SPSS” from the “Export” menu. The control file will have the extension CDP. When the program asks you if you want to export the corresponding RAW datafile, confirm the dialogue box.

When you are within SPSS, open this CDP file through the file menu → Open → other (not the data file). There are two program lines that need to be modified. The first one is toward the beginning part of the program and it reads:

```
data list file = "<DRIVE>:\<PATH>\<FILENAME>.RAW"
```

You should specify the appropriate drive, path, and filename. For example, if you are creating a SAV file for PBOOKLET within the same directory as the WinDEM data directory, then this line should be changed to your current setting, for example:

```
data list file = "C:\Program File\IEA\WinDEM\Data\pquest.RAW"
```

The second line to be modified is found at the end of the program:

```
save outfile = "<DRIVE>:\<PATH>\<FILENAME>.SAV" /COMPRESSED.
```

This line then should be modified to your current setting, for example:

```
save outfile = "C:\Program File\IEA\WinDEM\Data\pquest.SAV" /COMPRESSED.
```

For the alphanumerical variables, you will need to:

- add (a) after the variable name for the data list commend;
- delete variable names for the value label command;
- delete missing value commands;
- delete recode statements;
- change (fx.y) to (ax) in the format statements (where x represents a number of digits).

Now save this CDP file and select all the text to run. At the end of this procedure, your SPSS file PBOOKLET.SAV should be generated. Repeat this procedure for the other four files.