



SPSS Techniques Series: Pretest-Posttest Studies

SPSS

SPSS Techniques Series: Pretest-Posttest Studies

[Coding optical mark reader sheets](#)

[Describing the data](#)

[Choice of variable names](#)

[Scoring a test in SPSS](#)

[Analysis of Covariance](#)

Coding optical mark reader sheets

Pretest and posttest results can often be placed on scanning sheets. There are several ways to handle these.

First, if you are transferring the results from some other form to the scanning sheet, check to see if both the pretest and the posttest will fit on one scanning sheet. Suppose that you have 40 questions-the same questions on the pretest as on the posttest. Then you could record the pretest on one part of the scanning sheet, such as questions 1 to 40, and the post test on another part of the sheet, such as 51 to 90 (it is easier to have a gap between where the pretest stops and the post test begins, but not necessary).

Second, if the pretest and the posttest do not fit on one sheet, or if you had the individuals fill in their own sheets, then you will need 2 sheets for each person. It is absolutely essential to be able to match up the 2 sheets, in order to be able to control for subject differences when testing whether there has been a change from pre to post test. You will need to record an ID number on each survey, and a code indicating if the survey is pretest or posttest. If you have a control group as well as an experimental group, you will also need to have a code to indicated whether the person is part of the control group or the experimental group. Record all of this in the left hand part of the answer sheet. For example, say you have 10persons in your study. The first 5 were the control group, and the second 5 were the experimental group. Each person took a pretest and a posttest. Then you would have 20 answer sheets, and the codes on each sheet would be this:

first subject, pretest, control group:	001	1	1
first subject, posttest, control group:	001	2	1
second subject, pretest, control group:	002	1	1
second subject, posttest, control group:	002	2	1
third subject, pretest, control group:	003	1	1
third subject, posttest, control group:	003	2	1
fourth subject, pretest, control group:	004	1	1
fourth subject, posttest, control group:	004	2	1
fifth subject, pretest, control group:	005	1	1

fifth subject, posttest, control group:	005 2 1
sixth subject, pretest, experimental group:	006 1 2
sixth subject, posttest, experimental group:	006 2 2
seventh subject, pretest, experimental group:	007 1 2
seventh subject, posttest, experimental group:	007 2 2
eighth subject, pretest, experimental group:	008 1 2
eighth subject, posttest, experimental group:	008 2 2
ninth subject, pretest, experimental group:	009 1 2
ninth subject, posttest, experimental group:	009 2 2
tenth subject, pretest, experimental group:	010 1 2
tenth subject, posttest, experimental group:	010 2 2

Even if you have some sort of ID number such as a social security number, I still recommend putting a sequential ID number on the sheets, and also recording the same number on the survey. Sometimes a sheet will be entered out of order. When this happens, the pretest and the posttest sheets for an individual will not be next to each other in the file. The sequential ID number allows us to find the missing sheet quickly on the file and to move it to the correct spot. It is very difficult to look through social security numbers like this. If we record this same number on the survey, if at some point we question the data on the file we will be able to look up the original survey easily.

If you give the pretest and the posttest some distance of time apart, you may wish to enter the tests in two separate files. Then you can merge them together by the ID using SPSS. If you do this, it is very helpful to have the word "pretest" in the name of the file that contains the pretest, and "posttest" in the name of the file that contains the posttest. For example, you could use the file names "pretest.dat", and "posttest.dat".

Describing the data

1. When you have recorded both the pretest and the posttest on the same OMR sheet.

This is the easiest case. Here you don't have to worry about matching up sheets for the pre to the post test. In your data list statement, first define variables for the pretest questions, then for the post test questions.

Suppose you had a test with 20 questions. The answers to the pretest are recorded in questions 1 to 20 on the OMR sheet, and the answers to the post test are recorded in questions 51 to 70 on the same sheet. Say the file is called "THESIS.DAT". The data definition part of your program would be the following:

```
/* THESIS.SPS SPSSX program written by Mary Howard on 7/1/91.
/* This program demonstrates how to code a pre-post test situation
/* when the pre and post test are on the same OMR scan sheet.
data list
file= 'thesis.dat'
/id 51-53 group 57 sex 67(a) preq1 to preq20 70-89      postq1 to postq20 120-139
```

2. When you have two sheets for each person, sorted on the file with a pre and post test for each person.

```
/* THESIS.SPS SPSS program written by Mary Howard on 7/1/91.
/* This program demonstrates how to code a pre-post test situation
/* when the pre and post test are on two different OMR scan sheets,
/* consecutively in the file.
data list
file='thesis.dat' records=2
/1 id 51-53 group 57 sex 67(a) preq1 to preq20 70-89  /2 postq1 to postq20 70-89
```

3. When the pretest and the posttest are in two separate files.

```
/* THESIS.SPS SPSS program written by Mary Howard on 7/1/91.
/* This program demonstrates how to code a pre-post test situation
/* when the pre and post test are on two files.
data list
file='pretest.dat'
/id 60-62 sex 67(a) preq1 to preq20 70-89
file handle
tempfile
/name = 'pretest.systemfile'
sort cases by id
save
outfile=tempfile
execute
data list
file='posttest.dat'
/id 60-62 postq1 to postq20 70-89
sort cases by id
match files
file=tempfile
/file=*
/by idexecute
```

4. When the pretest and posttest are recorded on two separate sheets, and the sheets are in the same file, but they are not sorted.

When the sheets are in the file out of order, you must sort the sheets. Suppose you recorded an id of the sheet in the identification section of the scan sheet, followed by a code indicating the pre or post test, like this:

```
0001 1
0001 2
0002 10002 2
```

Assume this Id appears in columns 51 to 54 on the file, and the pre/post test code appears in column 56 of the file. A program to sort the test scoring sheets would be this:

```
/* fix.sps spss program written by Mary Howard
/* on 7/23/91. This program sorts a file where the scan sheets
/* are out of order.
data list
file='thesis.dat'
/pt1 1-50 (a) id 51-54 pt2 55 sheet 56 pt3 57-100(a)
pt4 101-200(a) pt5 201-300(a) pt6 301-400(a)
sort cases by id sheet
file handle
newfile
/name = 'thesis-2.dat'
write
outfile=newfile
/pt1, id, pt2, sheet, pt3,pt4,pt5,pt6execute
```

Once the data are in order, you can then proceed with the sorted file as is demonstrated in (2) above.

Choice of variable names

In fields such as psychology several instruments are often combined to form the entire questionnaire. Say the questionnaire is made of these instruments: the Beck scale, the Anxiety scale, and the Attitude scale. Let's say there are 40 questions on the Beck scale, 20 on the Anxiety scale, and 25 on the Attitude scale. Sometimes there are also scales that are given only on the pretest but not the posttest or vice versa, such as demographic questions. Variable names that include some indication of which scale the question is part of help enormously in later coding. If you are using two scan sheets to record the results, it also helps to record the pre and post results for each scale on the same location of the two sheets.

If we had recorded the Beck scale on questions 1 to 40 of the scan sheet, the Anxiety scale on questions 41 to 60 of the scan sheet, and the Attitude scale on questions 61 to 85. If we also had 10 demographic questions, given on the pretest only, we could record these on questions 86 to 95 of the pretest sheet only. Recall that scan sheet data starts in column 70 of the file. Then the first part of the SPSS program would be this:

```
/* mentalhealth.sps spss program written by Mary Howard on
/* 7/8/91. This demonstrates how to code a pre-post test situation
/* when the questionnaire is composed of several distinct
/* instruments.
data list
file='mentalhealth.dat' records =2
/1 id 51-52 group 53          /* 1=control, 2=experimental
prebck1 to prebck40 70-109
preanx1 to preanx20 110-129
preatt1 to preatt25 130-154
dem1 to dem10 155-164
/2 pstbck1 to pstbck40 70-109
pstanx1 to pstanx20 110-129          pstatt1 to pstatt25 130-154
```

Typically, subscales will be computed for each of the instruments in the questionnaire. Compute a subscale for both pre and post test, like this:

```
numeric prebck preanx preatt pstbck pstanx pstatt(f5.2)
compute prebck= mean(prebck1 to prebck40)
compute preanx= mean(preanx1 to preanx20)
compute preatt= mean(preatt1 to preatt25)
compute pstbck= mean(pstbck1 to pstbck40)
compute pstanx= mean(pstanx1 to pstanx20)compute pstatt= mean(pstatt1 to pstatt25)
```

You'll then want to see how the control and the experimental groups did. The DESCRIPTIVES procedure in SPSS will give means and standard deviations for each of the scales, and the T-TEST procedure can be used to compare the pre to the post test, like this:

```
title 'control only'
temporary
select if group=1
descriptives
variables= prebck pstbck preatt pstatt preanx pstanx
temporary
```

```

select if group=1
t-test
pairs=  prebck pstbck
/preatt pstatt
/preanx pstanx
title 'experimental only'
temporary
select if group=2
descriptives
variables = prebck pstbck preatt pstatt preanx pstanx
temporary
select if group=2
t-test
pairs = prebck pstbck
/preatt pstatt
/preanx pstanx
title ''

```

To compare the experimental group to the control group, compute the change from the pretest to the posttest for each scale, then do a grouped t-test:

```

numeric chgbck chgatt chganx(f5.2)
compute chgbck= pstbck-prebck
compute chgatt= pstatt-preatt
compute chganx= pstanx-preatt
title 'comparison of control to experimental group'
t-test
groups=group(1,2)
/variables = chbck chgatt chganx
title ''

```

Scoring a test in SPSS

If the instrument is a test, which you have given before and after some sort of experiment, then you'll need to score the test. One approach is to give 1 point for each correct answer, then add up the points.

First, determine which questions have which answers. Suppose for a 20 question test this is the key:

- Answer is "A" for questions 2,6,8,9,15 {"A" answers are coded as "1"}
- Answer is "B" for questions 1,3,10,16,20 {"B" answers are coded as "2"}
- Answer is "C" for questions 4,7,12 {"C" answers are coded as "3"}
- Answer is "D" for questions 5,11,13,14,17,18,19 {"D" answers are coded as "4"}

Assume also that the pretest and the posttest are identical. Create a point variable for each question in the data (20 pretest, 20 posttest). Assign each point variable a 1 if the answer to the question was correct; a 0 if the question was not correct. The code goes like this:

```

numeric prepts1 to prepts20 (f1)
numeric pstpts1 to pstpts20 (f1)
recode preq2 preq6 preq8 preq9 preq15
postq2 postq6 postq8 postq9 postq15

```

```

(1=1) (else=0) into
prepts2 prepts6 prepts8 prepts9 prepts15
pstpts2 pstpts6 pstpts8 pstpts9 pstpts15
recode preq1 preq3 preq10 preq16 preq20
postq1 postq3 postq10 postq16 postq20
(2=1) (else=0) into
prepts1 prepts3 prepts10 prepts16 prepts20
pstpts1 pstpts3 pstpts10 pstpts16 pstpts20
recode preq4 preq7 preq12
postq4 postq7 postq12
(3=1) (else=0) into
prepts4 prepts7 prepts12
pstpts4 pstpts7 pstpts12
recode preq5 preq11 preq13 preq14 preq17 preq18 preq19
postq5 postq11 postq13 postq14 postq17 postq18 postq19
(4=1) (else=0) into
prepts5 prepts11 prepts13 prepts17 prepts18 prepts19
pstpts5 pstpts11 pstpts13 pstpts17 pstpts18 pstpts19
numeric prescore pstscore (f2)
compute prescore= sum(prepts1 to prepts20)
compute pstscore= sum(pstpts1 to pstpts20)

```

Once you have computed the pretest score and the post test score, then you can compare the pretest to the posttest. To compare the pretest to the posttest for all subjects, do a paired t-test, like this:

```
t-test pairs= prescore pstscore
```

To select only the control or the experimental group, add "temporary", and a "select if" statements to the code:

```

title 'control only'
temporary
select if group=1
t-test
pairs=prescore pstscore
title 'experimental only'
temporary
select if group=2
t-test
pairs=prescore pstscore
title ''

```

If you want to determine if the experimental group improved more than the control group, first calculate the improvement, the post test score minus the pretest score. Then compare the improvement of the experimental group versus the improvement of the control group using a group t-test:

```

numeric improve(f2)
compute improve=pstscore-prescore
t-test
groups=group(1,2) /variables = prescore pstscore improve

```

Analysis of Covariance

Sometimes researchers will desire an Analysis of Covariance on a prepost study. This is the model: $pstscore = b_0 + b_1 \text{ group} + b_2 \text{ prescore} + b_3 \text{ group} * \text{prescore}$. This model tests whether we can predict the post score by the group (experimental or control), controlling for the pretest score of the individual (continuous). The code goes like this:

```
anova  
pstscore by group(1,2) with prescore /statistics=all
```

Alternatively, you can run this model using the Regression procedure. First you must code the group as 1,0 rather than 1,2. Also, an interaction term must be calculated prior to running the procedure. The code is this:

```
recode group(1=1) (2=0) /* 1 is experimental, 0 is control  
compute interac= group*prescore  
regression  
variables= prescore pstscore group interac  
/dependent= pstscore /method = enter
```

The results should be the same as that given by the analysis of covariance.

Source URL: <http://www.uni.edu/its/support/article/612>