

## [Essex Data Archive Study SN951]

Title: **Attitudes and Opinions of Senior Girls 1973**

Deposited by: **Survey Unit, Social Science Research Council**

File retrieval, restoration and additional documentation prepared Jan – Feb 2004 by:

John F Hall    La Noslière de Bas    50210 Notre Dame de Cenilly    France

Tel: (+) (33) (0) 2 33 45 91 47    email [johnfhall@orange.fr](mailto:johnfhall@orange.fr)

### **Background**

In Feb 1973, Sarah Abrams, daughter of the late Dr Mark Abrams (Director of the SSRC Survey Unit) was a pupil at St Paul's Girls' School, London, and editor of *Folio*, the school magazine run jointly with the boys' school. She wanted to write something for the magazine about the social and political opinions and attitudes of her fellow pupils, and mentioned this to her father, who arranged for her to come and meet himself and John Hall (Senior Research Fellow) at the Unit. Never able to resist the temptation of an interesting survey, Mark and John proposed (on methodological grounds) to use the idea to run a "quickie" survey at the school, in collaboration with Sarah, subject to consent of the High Mistress.

[Alternative para for public consumption:

In Feb 1973, a senior pupil at a girls' public school in the South of England, who was also the editor of her school magazine, wanted to write something for the magazine about the social and political opinions and attitudes of her fellow pupils. She approached the Survey Unit of the then Social Science Research Council for advice on a possible survey. A meeting was arranged with the late Dr Mark Abrams (Unit Director) and John Hall (Senior Research Fellow). Never able to resist the temptation of an interesting survey, Mark and John proposed (on methodological grounds, of course) to run a "quickie" survey at the school, subject to consent of the school authorities.]

John Hall made a quick trawl of available potential items, including some currently being used or considered for use in the Unit's Quality of Life surveys as part of its continuing programme of cross-national research in the development of subjective social indicators, and came up with a draft questionnaire (piloted on a few pupils for ease of understanding and use) followed by a final version which was administered to all girls in senior forms. Given the desperate timescale (end-of-term deadline for publication, less than a week for the fieldwork) the questionnaire was a real scissors and Sellotape effort, typed by John Hall at home on a manual typewriter with limited special character keys, with handdrawn lines and some handwritten elements, which he managed to get on to 4 sides of A4 (to get it on to 2 sheets double sided and keep reprographic costs down).

The questionnaire covered pocket money and expenditure, readership of newspapers and magazines, subjects taken or being taken at GCE 'O' and 'A' level (and level of interest these presented to girls) and influences on choice of such. Further questions covered perceived importance or not of selected objectives the school should have for its pupils and subjective rankings and personal expectations of future careers. A very limited amount of background information on pupils' families was followed by open-ended questions on perceived over- and under-privileged groups in society and by a selection of attitude scales gleaned from the literature or replicated from SSRC/SU's current work.

These included: (? modified versions of?):

New conservatism scale (Wilson-Patterson 35 items)

**Wilson, G.D.** (Ed.) *The Psychology of Conservatism*. London: Academic Press, 1973.

**Wilson, G.D.** & Patterson, J.R. A new measure of conservatism. *British Journal of Social and Clinical Psychology*, 1968, 7, 264-269.

Adolescent anti-change	(Himmelweit, 5 items)
Jaundiced view of life	(Himmelweit, 4 items)
Neuroticism	(Eysenck, 5 items)
Extraversion-Introversion	(Eysenck, 4 items)
Protest potential	(Inglehart, 15 items)

and a 21-item semantic differential scale to assess girls' perceptions of themselves

The editor distributed and collected the questionnaires and returned them to the Unit for coding and processing. Analysis was done on the ULCC computer via LSE and a **report\*\*** was prepared which served as the basis for an article in the school magazine.

### **The questionnaire**

A facsimile of the original questionnaire can be seen on file **0951.pdf**

There was no information on the questionnaire which could be used for data layout and data-preparation (it would have made for cluttered presentation and in any case there was no room!). The questionnaires were manually coded in-house by Eleanor Clutton-Brock and the data transferred to (?pre-printed?) coding sheets (can't remember unless there's an example extant), then punched on to 80-column Hollerith cards (3 per questionnaire).

**This makes it difficult to work direct from the questionnaire when performing data management and analysis, so if there is no data guide sheet, then one needs to be produced. Otherwise variable labels need to be checked to ensure the question number is included.**

Most of the questions were single response pre-coded on the questionnaire and these were single-punched on cards 1 and 2. Codes for some multiple response questions were multi-punched, but facilities for handling multi-punching of columns were not available in SPSS at the time, and so codes for multiple response questions on readership of newspapers and magazines were spread out and single-punched on card 3 for input to a very early version of SPSS. Some data seem not to be present ( Questions 6, 7 and 8: 'O' and 'A' level subjects taking/taken, and pupil's interest therein. This seems odd, but unless any other documentation comes to light, we must assume the data lost or the questions not coded in the first place.)

### **Restoration of files**

Although the final version of the SPSS saved file was submitted to Essex on a mag tape, this has not been preserved. A later version kept at the Polytechnic of North London seems to have suffered the same fate, as the tape archive available only goes back to 1986. This is a great pity, as SPSS commands **export** and **import** would have saved a great deal of time and tears. The author is not

yet completely *au fait* with the Windows version, but has managed to recreate a new saved file from the original setup files supplied by Essex.

\*\* Mark Abrams and John Hall **Attitudes and Opinions of Girls in Senior Forms**  
SSRC Survey Unit, March 1973 (mimeo 20pp)

[NB Author hasn't worked out how to do footnotes yet, or superscript characters]

Since 1973 there have been many subsequent releases of SSPS, not just for mainframe, but also for PC and most recently for Windows. The Windows release 11 has now got most, but not all, of the facilities of mainframe release 4. SPSS syntax has completely changed, and so many setup jobs simply will not work. Thus (with apologies to Ronald Searle) the file supplied defined data thus:

```
RUN NAME      TRINIANS CREATION
PAGE SIZE    NOEJECT
FILE NAME    TRINIANS
VARIABLE LIST FORM NUMBER MONEY YEARBORN MONTH VAR111 TO VAR119 JOB1 TO JOB5
              JOBAT25 SUCCESS1 SUCCESS2 LIKELY FATHER MOTHER PARENTS WEEKENDS
              SISTERS BROTHERS ELDEST VAR142 TO VAR176 VAR205 TO VAR234 VAR237
              TO VAR266 VAR270 TO VAR276 VAR305 TO VAR312 VAR314 VAR317 TO
              VAR339 VAR341 TO VAR349 VAR353 TO VAR364
INPUT MEDIUM INDATA
INPUT FORMAT FIXED(F1.0,F2.0,1X,F3.1,F1.0,F2.0,9F1.0,6A2,2A1,43F1.0/
              4X,30F1.0,2X,30F1.0,3X,7F1.0/
              4X,8F1.0,1X,F1.0,2X,23F1.0,1X,9F1.0,3X,12F1.0)
N OF CASES   216
```

But this had to be changed to:

```
data LIST records 3
      /1 FORM 1 NUMBER 2-3 MONEY 5-7 (1) YEARBORN 8 MONTH 9-10 VAR111 TO VAR119 11-19
      xOB1 TO xOB5 xOBAT25 20-31 (a)
      xUCCESS1 xUCCESS2 32-33 (a)
      LIKELY FATHER MOTHER PARENTS WEEKENDS
      SISTERS BROTHERS ELDEST VAR142 TO VAR176 34-76
      /2 VAR205 TO VAR234 5-34 VAR237 TO VAR266 37-66 VAR270 TO VAR276 70-76
      /3 VAR305 TO VAR312 5-12 VAR314 14 VAR317 TO VAR339 17-39
      VAR341 TO VAR349 41-49 VAR353 TO VAR364 53-64.
```

A second problem was trying to read the data from an external file. On my machine, SPSS could not find the data file specified, or did not like the way it was defined. Eventually it was quicker to copy the raw data file into the setup job and run it with **begin data** and **end data**. The eventual saved file was generated over several runs.

In the original version of SPSS it was possible to read in variables in alpha format and then **recode** them with a (**convert**) keeping the same variable names. This is no longer permitted as **string** variables (as they are now called) can only be converted into a new set of variables. Therefore the first letter of the initial variables to be read as strings was changed to x (eg **JOB1** was read in as **xOB1**) to create intermediate variables and a later **recode (convert)**ed them **into** the original names as specified in 1973; the intermediate variables were then deleted from the file.

This entailed modifications to the data transformation commands which were tedious rather than complicated. The variable labels and value labels needed modification to get rid of single primes and full stops, which took several runs as they were quite difficult to spot, but with the sheer speed of SPSS it was quicker to run jobs and look at the error reports, then delete the output file without saving it.

**SPSS still generates far too much output and could do with a facility for automatically keeping only two editions of output files, or at least having a prompt “Do you want to keep the output?” instead of clicking on the x and then answering a question.**

Also in 1973 there were no facilities for lower case letters or for automatic variable generation other than by **VARxxx TO VARyyy**. Later releases allowed names with any letter of the alphabet, but still only in capital letters (eg **Q1 to Q10**): nowadays lower case letters are allowed for names in setup jobs, but will be printed as capitals in output. **There is still no facility for generating names by e.g. Q1a to Q1g.**

The author has a distinct preference for operating via syntax files rather than ‘point and click’ on a menu, which horrifies him and is confusing and exasperating to use because not all the information needed is displayed in the view. Because at SSRC, and later at PNL, he and his colleagues were handling large numbers of surveys and even larger numbers of SPSS runs he developed a system for naming of files in which file names indicated what kind of run it was and file extensions what kind of file. Thus:

<b>TRINIANS.DAT</b>	contains raw data for the Trinians survey	
<b>TRINIANS.SPS</b>	would be a SPSS setup file generating output file	<b>TRINIANS.LST</b>
<b>TRINIANS.SYS</b>	would be the saved system file	
<b>TRINIANS.DOC</b>	would be a documentation file for the Trinians survey.	

. . and so on for **RECODE1.SPS RECODE2.SPS VARLABS.SPS VALLABS.SPS**

For a full explanation and of SSRC/SU and PNL/SRU conventions for variable naming, see file **NAMES.DOC**

**FREQ1.SPS** and **TAB1.SPS** generate **FREQ1.LST** and **TAB1.LST** (frequencies and tabulations) Even the extension names have been changed over the years, so even though **.sps** is the same, **.lst** became **.lis** and then **.spo**, whereas **.dat** now seems to indicate a WordPerfect file and **.doc** a file for MS-Word! Self-evidently jobs like **FACTOR.SPS** and **ANOVA.SPS** are easy to find in a directory and indicate the contents better than **SYNTAXddd** etc. At least two and sometimes three copies of all files were backed up on mag tape, and in cases where significant and substantial changes had been, there would be two or three previous editions of each file backed up as well.

SPSS for Windows doesn’t like this convention for names and extensions, but it doesn’t take long to learn to leave the extensions off and use the default SPSS (implied) extensions.

So far this restoration has taken 15 hours on 17 Jan and 5 hours on 18 Jan. and even more on subsequent days. The file has all the original variable and value labels in block capitals, except where some editing has been done. A first frequency count has thrown up some variables which have unexpected values or values with no labels, plus a few values still to be declared as missing. Also, the variable labels need to be checked to make sure the question numbers are included, as otherwise analysis would be a nightmare as the only documentation so far available is an unannotated questionnaire.

At least this now exists, but caused problems when printing from .pdf as the printer kept having a memory overflow and two of the pages wouldn’t fit properly, so even this is now a scissors and paste job!

[NB Should the relevant bits of the transformations and labels be included here (if I can find them all!) or as an appendix? Originals are on **d951.sps**, amendments (perhaps not all) on **syntax2.sps**]

JFH Sunday 18 January 2004 12:50 hrs

Tidied up missing values which, though declared seemed not to work and sorted value labels for some variables where full stop abbreviations made SPSS stop working. Like I said, tedious, but at least it's done. Current labelling very ugly and might have been quicker to retype the lot with decent lower case printing for output. File needs rearranging to get variables in a logical order, or at least questionnaire needs annotating by hand to indicate variable names.

Phase 1 complete at last!

JFH Sunday 18 January 2004 1500 hrs

Printed up some preliminary documentation last night from SPSS setup files and output from **data list** and **display**. There seem to be some variables missing, so need to check original data. Variables were not declared in questionnaire order for some (probably perfectly good) reason. Marked up copy of questionnaire with varnames and data locations. Some of these will need to be changed to conform to PNL-SRU conventions, and it would be useful to have at least rudimentary user manual with full question text, coding instructions, data locations and transformations, plus a frequency count (raw n only, but how to do this with SPSS **frequencies** which gives everything but the kitchen sink!)

JFH Monday 19 Jan

Tue 20 Jan. 04

Renamed variables from VARxxx and mnemonics to vddd (except derived vars) and reordered variables into order as entered. **This is not the same as the questionnaire**. Deleted superfluous and intermediate variables and added a couple of labels. Must find out how coding was done for Q2 Weeklies and others: also data for Q3 enjoyment of Folio.

Wed 21 Jan

Checked original data files to see what was coded where for multipunching. There is some, but apparently nothing for qq6-8. Printout of data file does not retain fixed width columns, so very difficult to read. Easier to use SPSS to write out a new data file. Our full conventions would have left a space after the serial number and a blank column somewhere in the middle of each card so that a printout will reveal codes that have slipped forwards or backwards (easily done when punching long lists of digits). This would be done separately for each card so that the blanks show up as a blank vertical column. Can't remember who did the spreading out, or where, but probably Jim Ring, who had by then joined SSRC/SU from LSE.

Thu 22 Jan. 04

Amendments to log of work done (confidentiality). Must really edit setup files to use lower case letters for labelling. If I could work out how to do it, the info on the data editor is enough to create a codebook key, but frequencies produces too much, if all we need is the raw codes and counts.

Fri 23 Jan. 04

Had a shot at multiple response tables, but SPSS won't do recodes into same vars, so had to create new vars for newspaper readership etc. Also Sundays and monthlies have been given labels in common, so needed to split these.

Being lazy, I've been trying to find quick ways of doing things, which is frustrating, but I'm learning my way round the editing facilities of SPSS and Word, and using whichever is quicker for me. So I find it's quicker to copy chunks of text out of SPSS setup files into Word, use that to change cases (usually whole file from upper to lower) and make mass substitutions to put some capitals back, then save as a .txt file. Latter can then be copied into a .sps file and run. Main problem is keeping track of all the changes and filenames, but am using old conventions of **varlab...** and **vallab....** for these plus **mult....** for multiple response setups. There's a lot of complex programming and trial and error in some of these, but there's no real need to include them in the main documentation except for SPSS buffs to show a few tricks of the trade.

The basic data set has multiple responses spread out as binary data in 1's and 0's, but for some applications the 1's need to be recoded to an ordinary coding sequence of 1 to n. In the former case tabulations can be done in binary format and the tables make sense, but only if the var label includes the code reference: in the latter, it is only necessary to put value labels on the first variable in the group = list even though this may seem bizarre to the novice user as all codes except the first one will not exist for the first variable. Question is whether to save the converted variables and labelling on the main file (eg by using **Mddd** instead of **Vddd** to indicate part of a set of variables for use in **mult response**

Hopefully have now managed to get file into presentable and usable format. One or two more mult response lists to sort out, but some base vars need checking first to see what's in there. Also the var sequence doesn't match the questionnaire sequence for precoded responses, but this may be due to in-house coding. Not sure who did this: could have been Sara herself or a trainee researcher, Eleanor Clutton-Brock.

To produce a multiple response frequency table in binary format..

```
mult response /group = Dailies 'Daily newspapers read'  
(v305 to v314 (1))  
/freq dailies.
```

```
Group DAILIES Daily newspapers read  
(Value tabulated = 1)
```

Dichotomy label	Name	Count	Pct of Responses	Pct of Cases
Q2 Daily papers - Express	V305	29	10.2	13.4
Q2 Daily papers - Mail	V306	23	8.1	10.6
Q2 Daily papers - Mirror	V307	5	1.8	2.3
Q2 Daily papers - Morning Star	V308	1	.4	.5
Q2 Daily papers - Sun	V309	1	.4	.5
Q2 Daily papers - Telegraph	V310	55	19.4	25.5
Q2 Daily papers - Times	V311	86	30.3	39.8
Q2 Daily papers - Guardian	V312	46	16.2	21.3
Q2 Daily papers - None read	V314	38	13.4	17.6
		-----	-----	-----
	Total responses	284	100.0	131.5

0 missing cases; 216 valid cases

But an attempt to produce the alternate format with...

```

recode v305 (1=1)/v306(1=2)/v307(1=3)/v308(1=4)/v309(1=5)/v310(1=6)/v311(1=7)/v312(1=8)/v314(1=0).
value labels v305 1 'Daily Express'
                  2 'Daily Mail'
                  3 'Daily Mirror'
                  4 'Morning Star'
                  5 'Sun'
                  6 'Daily Telegraph'
                  7 'Times'
                  8 'Guardian'
                  0 'None'.
mult response /group = Dailies 'Daily newspapers read'
                  (v305 to v314 (0,8))
          /freq dailies.

```

produces exactly the same table and so the following is needed....

```

do repeat
    x1=v305 to v314
    /x2=m305 to m312 m314.
compute x2 = x1.
end repeat.
recode m305 (1=1) /m306(1=2) /m307(1=3) /m308(1=4) /m309(1=5) /m310(1=6)
      /m311(1=7) /m312(1=8) /m314(1=0).
missing values m305 to m314 (0).
if v314=1 m314=9.
value labels m305 1 'Daily Express'
                  2 'Daily Mail'
                  3 'Daily Mirror'
                  4 'Morning Star'
                  5 'Sun'
                  6 'Daily Telegraph'
                  7 'Times'
                  8 'Guardian'
                  9 'None'.
mult response /group = Dailies 'Daily newspapers read'
                  (m305 to m314 (0,9)).

```

Group DAILIES Daily newspapers read

Category label	Code	Count	Pct of Responses	Pct of Cases
Daily Express	1	29	10.2	13.4
Daily Mail	2	23	8.1	10.6
Daily Mirror	3	5	1.8	2.3
Morning Star	4	1	.4	.5
Sun	5	1	.4	.5
Daily Telegraph	6	55	19.4	25.5
Times	7	86	30.3	39.8
Guardian	8	46	16.2	21.3
None	9	38	13.4	17.6
		-----	-----	-----
	Total responses	284	100.0	131.5

0 missing cases; 216 valid cases

The scales at the end need to be adjusted to give a true zero point, by subtracting the number of items in the scale from the score.