Technical details

In 2012 the sample for the British Social Attitudes survey was split into three equally-sized portions. Each portion was asked a different version of the questionnaire (versions A, B and C). Depending on the number of versions in which it was included, each 'module' of questions was thus asked either of the full sample (3,248 respondents) or of a random third or two-thirds of the sample. The structure of the questionnaire can be found at www.bsa-30.natcen.ac.uk.

Sample design

The British Social Attitudes survey is designed to yield a representative sample of adults aged 18 or over. Since 1993, the sampling frame for the survey has been the Postcode Address File (PAF), a list of addresses (or postal delivery points) compiled by the Post Office.[1]

For practical reasons, the sample is confined to those living in private households. People living in institutions (though not in private households at such institutions) are excluded, as are households whose addresses were not on the PAF.

The sampling method involved a multi-stage design, with three separate stages of selection.

Selection of sectors

At the first stage postcode sectors were selected systematically from a list of all postal sectors in Great Britain. Before selection, any sectors with fewer than 500 addresses were identified and grouped together with an adjacent sector; in Scotland all sectors north of the Caledonian Canal were excluded (because of the prohibitive costs of interviewing there). Sectors were then stratified on the basis of:

- 37 sub-regions;
- population density, (population in private households/area of the postal sector in hectares), with variable banding used in order to create three equal-sized strata per sub-region; and
- ranking by percentage of homes that were owner-occupied.

This resulted in the selection of 242 postcode sectors, with probability proportional to the number of addresses in each sector.

Selection of addresses

Twenty-eight addresses were selected in each of the 242 sectors or groups of sectors. The issued sample was therefore 242 x 28 = 6,776 addresses, selected by starting from a random point on the list of addresses for each sector, and choosing each address at a fixed interval. The fixed interval was calculated for each sector in order to generate the correct number of addresses.

The Multiple-Occupancy Indicator (MOI) available through PAF was used when selecting addresses in Scotland. The MOI shows the number of

accommodation spaces sharing one address. Thus, if the MOI indicated more than one accommodation space at a given address, the chances of the given address being selected from the list of addresses would increase so that it matched the total number of accommodation spaces. The MOI is largely irrelevant in England and Wales, as separate dwelling units (DUs) generally appear as separate entries on PAF. In Scotland, tenements with many flats tend to appear as one entry on PAF. However, even in Scotland, the vast majority (98.9 per cent) of MOIs in the sample had a value of one. The remainder had MOIs greater than one. The MOI affects the selection probability of the address, so it was necessary to incorporate an adjustment for this into the weighting procedures (described below).

Selection of individuals

Interviewers called at each address selected from PAF and listed all those eligible for inclusion in the British Social Attitudes sample – that is, all persons currently aged 18 or over and resident at the selected address. The interviewer then selected one respondent using a computer-generated random selection procedure. Where there were two or more DUs at the selected address, interviewers first had to select one DU using the same random procedure. They then followed the same procedure to select a person for interview within the selected DU.

Weighting

The weights for the British Social Attitudes survey correct for the unequal selection of addresses, DUs and individuals, and for biases caused by differential non-response. The different stages of the weighting scheme are outlined in detail below.

Selection weights

Selection weights are required because not all the units covered in the survey had the same probability of selection. The weighting reflects the relative selection probabilities of the individual at the three main stages of selection: address, DU and individual. First, because addresses in Scotland were selected using the MOI, weights were needed to compensate for the greater probability of an address with an MOI of more than one being selected, compared with an address with an MOI of one. (This stage was omitted for the English and Welsh data). Secondly, data were weighted to compensate for the fact that a DU at an address that contained a large number of DUs was less likely to be selected for inclusion in the survey than a DU at an address that contained fewer DUs. (We used this procedure because in most cases where the MOI is greater than one, the two stages will cancel each other out, resulting in more efficient weights.) Thirdly, data were weighted to compensate for the lower selection probabilities of adults living in large households, compared with those in small households.

At each stage the selection weights were trimmed to avoid a small number of very high or very low weights in the sample; such weights would inflate standard errors, reducing the precision of the survey estimates and causing the weighted sample to be less efficient. Less than one per cent of the selection weights were trimmed at each stage.

Non-response model

It is known that certain subgroups in the population are more likely to respond to surveys than others. These groups can end up over represented in the sample, which can bias the survey estimates. Where information is available about non-responding households, the response behaviour of the sample members can be modelled and the results used to generate a non-response weight. This non-response weight is intended to reduce bias in the sample resulting from differential response to the survey.

The data was modelled using logistic regression, with the dependent variable indicating whether or not the selected individual responded to the survey. Ineligible households[2] were not included in the non-response modelling. A number of area-level and interviewer observation variables were used to model response. Not all the variables examined were retained for the final model: variables not strongly related to a household's propensity to respond were dropped from the analysis.

The variables found to be related to response were: region, the relative condition of the immediate local area, the relative condition of the address, population density, and whether there were entry barriers to the selected address. The model shows that response increases if there are no barriers to entry (for instance, if there are no locked gates around the address and no entryphone) and if the general condition of the address is the same or better than other addresses in the area. Response decreases if the relative condition of the immediate surrounding area is mainly good or fair, and decreases as population density increases. Response is also higher for addresses in the North and in the Midlands. The full model is given in Table A.1.

Table A.1 The final non-response model

Variable	B	S.E.	Wald	Df	Sig.	Odds
Region			43.3	10	0.000	
North East	0.32	0.124	6.8	1	0.009	1.38
North West	0.25	0.091	7.8	1	0.005	1.29
Yorks. and Humber	-0.06	0.098	0.4	1	0.538	0.94
East Midlands	0.33	0.099	10.8	1	0.001	1.38
West Midlands	0.30	0.098	9.7	1	0.002	1.36
East of England	0.09	0.093	1.0	1	0.323	1.10
London	0.09	0.101	0.7	1	0.387	1.09
South East	0.01	0.088	0.0	1	0.926	1.01
South West	-0.03	0.097	0.1	1	0.739	0.97
Wales	0.08	0.118	0.5	1	0.472	1.09
Scotland	(baseline)					
Barriers to address						
No barriers	0.38	0.069	30.3	1	0.000	1.46
One or more	(baseline)					
Relative condition of the local area			32.7	2	0.000	
Mainly good	-0.22	0.043	26.4	1	0.000	0.80
Mainly fair	-0.42	0.115	13.2	1	0.000	0.66
Mainly bad or very bad	(baseline)					
Relative condition of the address			35.7	2	0.000	
Better	0.52	0.110	22.6	1	0.000	1.69
About the same	0.05	0.083	0.4	1	0.528	1.05
Worse	(baseline)					
Population density (population in private households/ area of postcode sector in hectares)	0.00	0.001	24.6	1	0.000	0.99
Constant	0.06	0.123	.2	1	0.650	1.06

The response is 1 = individual responding to the survey, 0 = non-response
Only variables that are significant at the 0.05 level are included in the model
The model R2 is 0.023 (Cox and Snell)
B** is the estimate coefficient with standard error **S.E.

*The **Wald**-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom (**df**). If the test is significant (**sig.** < 0.05), then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model*

The non-response weight was calculated as the inverse of the predicted response probabilities saved from the logistic regression model. The non-response weight was then combined with the selection weights to create the final non-response weight. The top one per cent of the weight were trimmed before the weight was scaled to the achieved sample size (resulting in the weight being standardised around an average of one).

Calibration weighting

The final stage of weighting was to adjust the final non-response weight so that the weighted sample matched the population in terms of age, sex and region.

Table A.2 Weighted and unweighted sample distribution, by region, age and sex

	Population	Unweighted respondents	Respondent weighted by selection weight only	Respondent weighted by un-calibrated non-response weight	Respondent weighted by final weight
Region	%	%	%	%	%
North East	4.3	4.6	4.5	4.1	4.3
North West	11.5	13.1	13.3	12.5	11.5
Yorks. and Humber	8.6	8.3	8.0	8.4	8.6
East Midlands	7.4	9.4	9.6	8.7	7.4
West Midlands	9.0	9.8	9.7	9.0	9.0
East of England	9.5	10.3	10.5	10.3	9.5
London	13.2	9.2	10.0	11.5	13.2
South East	14.0	12.5	12.9	13.3	14.0
South West	8.7	8.9	8.3	8.7	8.7
Wales	5.0	4.9	4.7	4.6	5.0
Scotland	8.8	9.1	8.5	8.9	8.8
Age and sex	%	%	%	%	%
M 18–24	5.9	2.4	3.4	3.4	5.9
M 25–34	8.5	4.9	5.5	5.7	8.5
M 35–44	9.0	8.0	8.3	8.4	9.0
M 45–54	8.9	7.8	8.4	8.4	8.9
M 55–59	3.7	3.5	3.3	3.2	3.7
M 60–64	3.9	4.5	4.5	4.4	3.9
M 65+	11.7	13.3	12.8	12.4	11.7
F 18–24	6.0	4.4	5.3	5.5	6.0
F 25–34	8.5	7.9	7.8	8.2	8.5
F 35–44	8.8	9.8	9.2	9.3	8.8
F 45–54	8.7	8.5	9.5	9.4	8.7
F 55–59	3.6	3.9	4.0	4.0	3.6
F 60–64	3.8	4.4	4.3	4.2	3.8
F 65+	9.3	16.6	13.8	13.5	9.3
Base	*48,306,543*	*3248*	*3248*	*3248*	*3248*

Only adults aged 18 or over are eligible to take part in the survey, therefore the data have been weighted to the British population aged 18+ based on 2011 Census data from the Office for National Statistics/General Register Office for Scotland.

The survey data were weighted to the marginal age/sex and region distributions using raking-ratio (or rim) weighting. As a result, the weighted data should exactly match the population across these three dimensions. This is shown in Table A.2.

The calibration weight is the final non-response weight to be used in the analysis of the 2012 survey; this weight has been scaled to the responding sample size. The range of the weights is given in Table A.3.

Table A.3 Range of weights

	N	Minimum	Mean	Maximum
DU and person selection weight	3248	0.55	1.00	2.22
Un-calibrated non-response weight	3248	0.42	1.00	2.52
Final calibrated non-response weight	3248	0.26	1.00	4.77

Effective sample size

The effect of the sample design on the precision of survey estimates is indicated by the effective sample size (neff). The effective sample size measures the size of an (unweighted) simple random sample that would achieve the same precision (standard error) as the design being implemented. If the effective sample size is close to the actual sample size, then we have an efficient design with a good level of precision. The lower the effective sample size is, the lower the level of precision. The efficiency of a sample is given by the ratio of the effective sample size to the actual sample size. Samples that select one person per household tend to have lower efficiency than samples that select all household members. The final calibrated non-response weights have an effective sample size (neff) of 2,446 and efficiency of 75 per cent.

All the percentages presented in this report are based on weighted data.

Questionnaire versions

Each address in each sector (sampling point) was allocated to one of the portions of the sample: A, B or C. As mentioned earlier, a different version of the questionnaire was used with each of the three sample portions. If one serial number was version A, the next was version B and the third version C. Thus, each interviewer was allocated 10 cases from each of versions A, B and C. There were 2,259 issued addresses for versions A and B, and 2,258 for version C.

Fieldwork

Interviewing was mainly carried out between June and September 2012, with a small number of interviews taking place in October and November.

Fieldwork was conducted by interviewers drawn from NatCen Social Research's regular panel and conducted using face-to-face computer-assisted interviewing[3] Interviewers attended a one-day briefing conference to familiarise them with the selection procedures and questionnaires, with the exception of very experienced interviewers who completed a self-briefing containing updates to the questionnaire and procedures.

The mean interview length was 61 minutes for version A of the questionnaire, 63 minutes for version B and 62 minutes for version C.[4] Interviewers achieved an overall response rate of between 53.2 and 53.5 per cent. Details are shown in Table A.4.

Table A.4 Response rate[1] on British Social Attitudes, 2012

	Number	Lower limit of response (%)	Upper limit of response (%)
Addresses issued	6776		
Out of scope	674		
Upper limit of eligible cases	6102	100.0	
Uncertain eligibility	32	0.5	
Lower limit of eligible cases	6070		100.0
Interview achieved	3248	53.2	53.5
With self-completion	2866	47.0	47.2
Interview not achieved	2822	46.2	46.5
Refused[2]	2133	35.0	35.1
Non-contacted[3]	319	5.2	5.3
Other non-response	370	6.1	6.1

1 Response is calculated as a range from a lower limit where all unknown eligibility cases (for example, address inaccessible, or unknown whether address is residential) are assumed to be eligible and therefore included in the unproductive outcomes, to an upper limit where all these cases are assumed to be ineligible and therefore excluded from the response calculation

2 'Refused' comprises refusals before selection of an individual at the address, refusals to the office, refusal by the selected person, 'proxy' refusals (on behalf of the selected respondent) and broken appointments after which the selected person could not be recontacted

3 'Non-contacted' comprises households where no one was contacted and those where the selected person could not be contacted

As in earlier rounds of the series, the respondent was asked to fill in a self-completion questionnaire which, whenever possible, was collected by the interviewer. Otherwise, the respondent was asked to post it to NatCen Social Research. If necessary, up to three postal reminders were sent to obtain the self-completion supplement.

A total of 382 respondents (12 per cent of those interviewed) did not return their self-completion questionnaire. Version A of the self-completion questionnaire was returned by 86 per cent of respondents to the face-to-face interview, version B of the questionnaire was returned by 90 per cent and version C by 89 per cent. As in previous rounds, we judged that it was not necessary to apply additional weights to correct for non-response to the self-completion questionnaire.

Advance letter

Interviewers were supplied with letters describing the purpose of the survey and the coverage of the questionnaire, which they posted to sampled addresses before making any calls.[5]

Analysis variables

A number of standard analyses have been used in the tables that appear in this report. The analysis groups requiring further definition are set out below. For further details see Stafford and Thomson (2006). Where there are references to specific question numbers, the full question text, including frequencies, can be found at www.bsa-30.natcen.ac.uk.

Region

The dataset is classified by 12 regions, formerly the Government Office Regions.

Standard Occupational Classification

Respondents are classified according to their own occupation, not that of the 'head of household'. Each respondent was asked about their current or last job, so that all respondents except those who had never worked were coded. Additionally, all job details were collected for all spouses and partners in work.

Since the 2011 survey, we have coded occupation to the Standard Occupational Classification 2010 (SOC 2010) instead of the Standard Occupational Classification 2000 (SOC 2000). The main socio-economic grouping based on SOC 2010 is the National Statistics Socio-Economic Classification (NS-SEC). However, to maintain time series, some analysis has continued to use the older schemes based on SOC 90 – Registrar General's Social Class and Socio-Economic Group – though these are now derived from SOC 2000 (which is derived from SOC 2010).

National Statistics Socio-Economic Classification (NS-SEC)

The combination of SOC 2010 and employment status for current or last job generates the following NS-SEC analytic classes:

- Employers in large organisations, higher managerial and professional
- Lower professional and managerial; higher technical and supervisory
- Intermediate occupations
- Small employers and own account workers
- Lower supervisory and technical occupations
- Semi-routine occupations
- Routine occupations

The remaining respondents are grouped as "never had a job" or "not classifiable". For some analyses, it may be more appropriate to classify respondents according to their current socio-economic status, which takes into account only their present economic position. In this case, in addition to the seven classes listed above, the remaining respondents not currently in paid work fall into one of the following categories: "not classifiable", "retired", "looking after the home", "unemployed" or "others not in paid occupations".

Registrar General's Social Class

As with NS-SEC, each respondent's social class is based on his or her current or last occupation. The combination of SOC 90 with employment status for current or last job generates the following six social classes:

I	Professional etc. occupations	
II	Managerial and technical occupations	'Non-manual'
III (Non-manual)	Skilled occupations	
III (Manual)	Skilled occupations	
IV	Partly skilled occupations	'Manual'
V	Unskilled occupations	

They are usually collapsed into four groups: I & II, III Non-manual, III Manual, and IV & V.

Socio-Economic Group

As with NS-SEC, each respondent's Socio-Economic Group (SEG) is based on his or her current or last occupation. SEG aims to bring together people with jobs of similar social and economic status, and is derived from a combination of employment status and occupation. The full SEG classification identifies 18 categories, but these are usually condensed into six groups:

- Professionals, employers and managers
- Intermediate non-manual workers
- Junior non-manual workers
- Skilled manual workers
- Semi-skilled manual workers
- Unskilled manual workers

As with NS-SEC, the remaining respondents are grouped as "never had a job" or "not classifiable".

Industry

All respondents whose occupation could be coded were allocated a Standard Industrial Classification 2007 (SIC 07). Two-digit class codes are used. As with social class, SIC may be generated on the basis of the respondent's current occupation only, or on his or her most recently classifiable occupation.

Party identification

Respondents can be classified as identifying with a particular political party on one of three counts: if they consider themselves supporters of that party, closer to it than to others, or more likely to support it in the event of a general election. The three groups are generally described respectively as *partisans*, *sympathisers* and *residual identifiers*. In combination, the three groups are referred to as 'identifiers'. Responses are derived from the following questions:

> *Generally speaking, do you think of yourself as a supporter of any one political party? [Yes/No]*
>
> *[If "No"/"Don't know"]*
> *Do you think of yourself as a little closer to one political party than to the others? [Yes/No]*
>
> *[If "Yes" at either question or "No"/"Don't know" at 2nd question]*
> *Which one?/If there were a general election tomorrow, which political party do you think you would be most likely to support?*
>
> *[Conservative; Labour; Liberal Democrat; Scottish National Party; Plaid Cymru; Green Party; UK Independence Party (UKIP)/Veritas; British National Party (BNP)/National Front; RESPECT/Scottish Socialist Party (SSP)/Socialist Party; Other party; Other answer; None; Refused to say]*

Income

Two variables classify the respondent's earnings [REarn] and household income [HHInc] (see www.bsa-30.natcen.ac.uk). The bandings used are designed to be representative of those that exist in Britain and are taken from the Family Resources Survey (see http://research.dwp.gov.uk/asd/frs/). Four derived variables give income deciles/quartiles: [RearnD], [REarnQ], [HHIncD] and [HHIncQ]. Deciles and quartiles are calculated based on individual earnings and household incomes in Britain as a whole.

Attitude scales

Since 1986, the British Social Attitudes surveys have included two attitude scales which aim to measure where respondents stand on certain underlying value dimensions – left–right and libertarian–authoritarian.[6] Since 1987 (except in 1990), a similar scale on 'welfarism' has also been included. Some of the items in the welfarism scale were changed in 2000–2001. The current version of the scale is shown below.

A useful way of summarising the information from a number of questions of this sort is to construct an additive index (Spector, 1992; DeVellis, 2003). This approach rests on the assumption that there is an underlying – 'latent' – attitudinal dimension which characterises the answers to all the questions within each scale. If so, scores on the index are likely to be a more reliable indication of the underlying attitude than the answers to any one question.

Each of these scales consists of a number of statements to which the respondent is invited to "agree strongly", "agree", "neither agree nor disagree", "disagree" or "disagree strongly".

The items are:

Left–right scale

Government should redistribute income from the better off to those who are less well off. [Redistrb]

Big business benefits owners at the expense of workers. [BigBusnN]

Ordinary working people do not get their fair share of the nation's wealth. [Wealth] [7]

There is one law for the rich and one for the poor. [RichLaw]

Management will always try to get the better of employees if it gets the chance. [Indust4]

Libertarian–authoritarian scale

Young people today don't have enough respect for traditional British values. [TradVals]

People who break the law should be given stiffer sentences. [StifSent]

For some crimes, the death penalty is the most appropriate sentence. [DeathApp]

Schools should teach children to obey authority. [Obey]

The law should always be obeyed, even if a particular law is wrong. [WrongLaw]

Censorship of films and magazines is necessary to uphold moral standards. [Censor]

Welfarism scale

The welfare state encourages people to stop helping each other. [WelfHelp]

The government should spend more money on welfare benefits for the poor, even if it leads to higher taxes. [MoreWelf]

Around here, most unemployed people could find a job if they really wanted one. [UnempJob]

Many people who get social security don't really deserve any help. [SocHelp]

Most people on the dole are fiddling in one way or another. [DoleFidl]

If welfare benefits weren't so generous, people would learn to stand on their own two feet. [WelfFeet]

Cutting welfare benefits would damage too many people's lives. [DamLives]

The creation of the welfare state is one of Britain's proudest achievements. [ProudWlf]

The indices for the three scales are formed by scoring the leftmost, most libertarian or most pro-welfare position, as 1 and the rightmost, most authoritarian or most anti-welfarist position, as 5. The "neither agree nor disagree" option is scored as 3. The scores to all the questions in each scale are added and then divided by the number of items in the scale, giving indices ranging from 1 (leftmost, most libertarian, most pro-welfare) to 5 (rightmost, most authoritarian, most anti-welfare). The scores on the three indices have been placed on the dataset.[8]

The scales have been tested for reliability (as measured by Cronbach's alpha). The Cronbach's alpha (unstandardised items) for the scales in 2012 are 0.82 for the left–right scale, 0.81 for the welfarism scale and 0.73 for the libertarian–authoritarian scale. This level of reliability can be considered "good" for the left–right and libertarian–authoritarian scales and "respectable" for the welfarism scale (DeVellis, 2003: 95–96).

Other analysis variables

These are taken directly from the questionnaire and to that extent are self-explanatory (see www.bsa-30.natcen.ac.uk). The principal ones are:

- Sex (Q. 63)
- Age (Q. 64)
- Household income (Q. 1020)
- Economic position (Q. 465)
- Religion (Q. 816)
- Highest educational qualification obtained (Q. 948)
- Marital status (Qs. 162–168)
- Benefits received (Qs. 543–598)

Sampling errors

No sample precisely reflects the characteristics of the population it represents, because of both sampling and non-sampling errors. If a sample were designed as a random sample (if every adult had an equal and independent chance of inclusion in the sample), then we could calculate the sampling error of any percentage, p, using the formula:

$$\text{s.e. (p)} = \sqrt{\frac{p(100 - p)}{n}}$$

where *n* is the number of respondents on which the percentage is based. Once the sampling error had been calculated, it would be a straightforward exercise to calculate a confidence interval for the true population percentage. For example, a 95 per cent confidence interval would be given by the formula:

$$p \pm 1.96 \times \text{s.e. (p)}$$

Clearly, for a simple random sample (srs), the sampling error depends only on the values of p and n. However, simple random sampling is almost never used in practice, because of its inefficiency in terms of time and cost.

As noted above, the British Social Attitudes sample, like that drawn for most large-scale surveys, was clustered according to a stratified multi-stage design into 242 postcode sectors (or combinations of sectors). With a complex design like this, the sampling error of a percentage giving a particular response is not simply a function of the number of respondents in the sample and the size of the percentage; it also depends on how that percentage response is spread within and between sample points.

The complex design may be assessed relative to simple random sampling by calculating a range of design factors (DEFTs) associated with it, where:

$$\text{DEFT} = \sqrt{\frac{\text{Variance of estimator with complex design, sample size n}}{\text{Variance of estimator with srs design, sample size n}}}$$

and represents the multiplying factor to be applied to the simple random sampling error to produce its complex equivalent. A design factor of one means that the complex sample has achieved the same precision as a simple random sample of the same size. A design factor greater than one means the complex sample is less precise than its simple random sample equivalent. If the DEFT for a particular characteristic is known, a 95 per cent confidence interval for a percentage may be calculated using the formula:

$$p \pm 1.96 \times \text{complex sampling error (p)}$$

$$= p \pm 1.96 \times \text{DEFT} \times \sqrt{\frac{p(100 - p)}{n}}$$

Table A.5 gives examples of the confidence intervals and DEFTs calculated for a range of different questions. Most background questions were asked of the whole sample, whereas many attitudinal questions were asked only of a third or two-thirds of the sample; some were asked on the interview questionnaire and some on the self-completion supplement.

Table A.5 Complex standard errors and confidence intervals of selected variables

	% (p)	Complex standard error of p	95% confidence interval	DEFT	*Base*
Classification variables					
Q. 274–279 Party identification (full sample)					
Conservative	27.0	1.0	25.0–29.1	1.291	*3248*
Labour	35.9	1.1	33.7–38.1	1.271	*3248*
Liberal Democrat	6.3	0.6	5.3–7.5	1.275	*3248*
Q. 809 Housing tenure (full sample)					
Owns	67.6	1.3	64.9–70.2	1.635	*3248*
Rents from local authority	9.9	0.8	8.4–11.6	1.518	*3248*
Rents privately/HA	21.3	1.1	19.2–23.5	1.536	*3248*
Q. 816 Religion (full sample)					
No religion	47.9	1.2	45.5–50.4	1.407	*3248*
Church of England	20.3	0.8	18.7–21.9	1.172	*3248*
Roman Catholic	8.7	0.5	7.7–9.8	1.098	*3248*
Q. 880 Age of completing continuous full-time education (full sample)					
16 or under	47.9	1.2	45.5–50.4	1.413	*3248*
17 or 18	20.3	0.9	18.7–22.1	1.244	*3248*
19 or over	27.6	1.1	25.5–29.8	1.386	*3248*
Q. 267 Home internet access (full sample)					
Yes	84.2	0.8	82.6–85.6	1.203	*3248*
No	15.8	0.8	14.3–17.4	1.202	*3248*
Q. 813 Urban or rural residence (full sample)					
A big city	10.1	1.2	7.9–12.8	2.308	*3248*
The suburbs or outskirts of a big city	27.7	1.9	24.0–31.6	2.452	*3248*
A small city/town	40.4	2.3	35.9–44.9	2.645	*3248*
Country village	18.1	1.7	15.0–21.7	2.499	*3248*
Farm/home in the country	3.3	0.6	2.3–4.7	1.851	*3248*
Attitudinal variables (face-to-face interview)					
Q. 375 Benefits for the unemployed are … (full sample)					
… too low	22.4	0.9	20.6–24.3	1.298	*3248*
… too high	51.3	1.2	48.8–53.7	1.422	*3248*
Q. 411 How serious a problem is traffic congestion in towns, cities (full sample)					
A very serious problem	10.6	0.7	9.3–12.2	1.348	*3248*
A serious problem	28.0	1.0	26.0–30.0	1.286	*3248*
Not a very serious problem	41.2	1.1	39.1–43.4	1.261	*3248*
Not a problem at all	20.0	1.0	18.1–22.1	1.446	*3248*

Table A.5 Complex standard errors and confidence intervals of selected variables (continued)

	% (p)	Complex standard error of p	95% confidence interval	DEFT	*Base*
Q. 317 How much child poverty in Britain today (full sample)					
There is no child poverty in Britain today	3.1	0.4	2.4–3.9	1.260	*3248*
There is very little child poverty in Britain today	12.7	0.7	11.3–14.2	1.252	*3248*
There is some child poverty in Britain today	43.0	1.0	41.0–45.0	1.180	*3248*
There is quite a lot of child poverty in Britain today	39.2	1.0	37.3–41.3	1.179	*3248*
Q. 443 Who should pay towards the cost of tuition (half sample)					
All students or their families should pay the	11.4	0.8	9.9–13.1	1.164	*2145*
Some students or their families should pay	67.6	1.2	65.3–70.0	1.178	*2145*
No students or their families should pay	19.7	0.9	17.9–21.6	1.089	*2145*
Attitudinal variables (self-completion)					
A48a B14a C13a Government should redistribute income from the better off to those who are less well off (full sample)					
Agree strongly	10.7	0.6	9.6–11.9	.998	*2855*
Agree	30.5	1.1	28.4–32.6	1.249	*2855*
Neither agree nor disagree	26.3	0.9	24.7–28.0	1.038	*2855*
Disagree	23.6	0.9	22.0–25.4	1.080	*2855*
Disagree strongly	6.4	0.6	5.4–7.7	1.260	*2855*
B10f Are the police well run or not well run (1/3 sample)					
Very well run	9.7	1.1	7.7–12.2	1.171	*956*
Well run	55.5	1.8	51.8–59.1	1.147	*956*
Not very well run	22.2	1.8	18.9–25.8	1.314	*956*
Not at all well run	4.0	1.1	2.8–5.6	1.078	*956*
A3a Do you think that women should work outside the home full-time, part-time or not at all ...when there is a child under school age (1/3 sample)					
Work full-time	5.4	0.8	4.0–7.2	1.082	*950*
Work part-time	43.2	2.0	39.2–47.2	1.268	*950*
Stay at home	33.2	1.9	29.6–37.1	1.251	*950*
A38a People should be able to travel by plane as much as they like (1/3 sample)					
Agree	64.9	2.0	61.0–68.7	1.262	*950*
Neither agree nor disagree	20.2	1.5	17.4–23.5	1.184	*950*
Disagree	8.7	1.1	6.7–11.1	1.208	*950*

Table A.5 shows that most of the questions asked of all sample members have a confidence interval of around plus or minus two to three per cent of the survey percentage. This means that we can be 95 per cent certain that the true population percentage is within two to three per cent (in either direction) of the percentage we report.

Variables with much larger variation are, as might be expected, those closely related to the geographic location of the respondent (for example, whether they live in a big city, a small town or a village). Here, the variation may be as large as six or seven per cent either way around the percentage found on the survey. Consequently, the design effects calculated for these variables in a clustered sample will be greater than the design effects calculated for variables less strongly associated with area. Also, sampling errors for percentages based only on respondents to just one of the versions of the questionnaire, or on subgroups within the sample, are larger than they would have been had the questions been asked of everyone.

Analysis techniques

Cohort analysis

A number of the chapters in this report employ 'pseudo-cohort' analysis (which we describe as 'cohort analysis' throughout). When applied to attitudinal research cohort analysis traditionally uses longitudinal survey data to examine how the views of the same individuals have changed over time. Pseudo-cohort analysis uses cross-sectional data and is based on the assumption that a particular age group within a given year is equivalent to an age group five years older, five years later.

Cohort analysis is used to explore whether change over time can be explained by generation, period or lifecycle effects – or a combination of the three:

- A **generational effect** can be identified when each successive generation expresses an attitude that is different to the one which preceded it. As a result, and when these differences all occur in a similar direction, change at the population level can be driven by the ageing of the population, as older generations die out and younger generations enter the population (in the case of British Social Attitudes, those aged 18+).
- A **period effect** can be identified when the views of all or most generations change in a consistent way within a particular period. This can often be linked to an external event.
- A **lifecycle effect** can be identified when the views of all of most generations change in a particular way during a particular life-stage such as adolescence or retirement or, slternatively, across the life-cycle.

While a cohort or generation is a subjective construct and these can be defined in a number of ways, in this report we have consistently allocated respondents to cohorts based on their decade of birth (for instance, those born between 1980 and 1989 are defined as the '1980s' cohort). This approach was adopted to explore in considerable detail how the attitudes of generations have changed over time and in relation to one another.

Data has only been included in the charts and tables produced to illustrate cohort analysis when data on the measure of interest is available for at least 100 cases in a given year. This means that, in some cases, data is not presented for the oldest cohort, given the small sample sizes involved.

Regression

Regression analysis aims to summarise the relationship between a 'dependent' variable and one or more 'independent' variables. It shows how well we can

estimate a respondent's score on the dependent variable from knowledge of their scores on the independent variables. It is often undertaken to support a claim that the phenomena measured by the independent variables cause the phenomenon measured by the dependent variable. However, the causal ordering, if any, between the variables cannot be verified or falsified by the technique. Causality can only be inferred through special experimental designs or through assumptions made by the analyst.

All regression analysis assumes that the relationship between the dependent and each of the independent variables takes a particular form. In linear regression it is assumed that the relationship can be adequately summarised by a straight line. This means that a one percentage point increase in the value of an independent variable is assumed to have the same impact on the value of the dependent variable on average, irrespective of the previous values of those variables.

Strictly speaking the technique assumes that both the dependent and the independent variables are measured on an interval-level scale, although it may sometimes still be applied even where this is not the case. For example, one can use an ordinal variable (e.g. a Likert scale) as a *dependent* variable if one is willing to assume that there is an underlying interval-level scale and the difference between the observed ordinal scale and the underlying interval scale is due to random measurement error. Often the answers to a number of Likert-type questions are averaged to give a dependent variable that is more like a continuous variable. Categorical or nominal data can be used as *independent* variables by converting them into dummy or binary variables; these are variables where the only valid scores are 0 and 1, with 1 signifying membership of a particular category and 0 otherwise.

The assumptions of linear regression cause particular difficulties where the *dependent* variable is binary. The assumption that the relationship between the dependent and the independent variables is a straight line means that it can produce estimated values for the dependent variable of less than 0 or greater than 1. In this case it may be more appropriate to assume that the relationship between the dependent and the independent variables takes the form of an S-curve, where the impact on the dependent variable of a one-point increase in an independent variable becomes progressively less the closer the value of the dependent variable approaches 0 or 1. *Logistic* regression is an alternative form of regression which fits such an S-curve rather than a straight line. The technique can also be adapted to analyse multinomial non-interval-level dependent variables, that is, variables which classify respondents into more than two categories.

The two statistical scores most commonly reported from the results of regression analyses are:

A measure of variance explained: This summarises how well all the independent variables combined can account for the variation in respondents' scores in the dependent variable. The higher the measure, the more accurately we are able in general to estimate the correct value of each respondent's score on the dependent variable from knowledge of their scores on the independent variables.

A parameter estimate: This shows how much the dependent variable will change on average, given a one-unit change in the independent variable (while holding all other independent variables in the model constant). The parameter estimate

has a positive sign if an increase in the value of the independent variable results in an increase in the value of the dependent variable. It has a negative sign if an increase in the value of the independent variable results in a decrease in the value of the dependent variable. If the parameter estimates are standardised, it is possible to compare the relative impact of different independent variables; those variables with the largest standardised estimates can be said to have the biggest impact on the value of the dependent variable.

Regression also tests for the statistical significance of parameter estimates. A parameter estimate is said to be significant at the five per cent level if the range of the values encompassed by its 95 per cent confidence interval (see also section on sampling errors) are either all positive or all negative. This means that there is less than a five per cent chance that the association we have found between the dependent variable and the independent variable is simply the result of sampling error and does not reflect a relationship that actually exists in the general population.

Factor analysis

Factor analysis is a statistical technique which aims to identify whether there are one or more apparent sources of commonality to the answers given by respondents to a set of questions. It ascertains the smallest number of *factors* (or dimensions) which can most economically summarise all of the variation found in the set of questions being analysed. Factors are established where respondents who gave a particular answer to one question in the set tended to give the same answer as each other to one or more of the other questions in the set. The technique is most useful when a relatively small number of factors are able to account for a relatively large proportion of the variance in all of the questions in the set.

The technique produces a *factor loading* for each question (or variable) on each factor. Where questions have a high loading on the same factor, then it will be the case that respondents who gave a particular answer to one of these questions tended to give a similar answer to each other at the other questions. The technique is most commonly used in attitudinal research to try to identify the underlying ideological dimensions which apparently structure attitudes towards the subject in question.

International Social Survey Programme

The International Social Survey Programme (ISSP) is run by a group of research organisations in different countries, each of which undertakes to field annually an agreed module of questions on a chosen topic area. Since 1985, an International Social Survey Programme module has been included in one of the British Social Attitudes self-completion questionnaires. Each module is chosen for repetition at intervals to allow comparisons both between countries (membership is currently standing at 48) and over time. In 2012, the chosen subject was Family, Work and Gender Roles, and the module was carried on version A of the self-completion questionnaire (Qs. 1a–33).[9]

Notes

1. Until 1991 all British Social Attitudes samples were drawn from the Electoral Register (ER). However, following concern that this sampling frame might be deficient in its coverage of certain population subgroups, a 'splicing' experiment was conducted in 1991. We are grateful to the Market Research Development Fund for contributing towards the costs of this experiment. Its purpose was to investigate whether a switch to PAF would disrupt the time series – for instance, by lowering response rates or affecting the distribution of responses to particular questions. In the event, it was concluded that the change from ER to PAF was unlikely to affect time trends in any noticeable ways, and that no adjustment factors were necessary. Since significant differences in efficiency exist between PAF and ER, and because we considered it untenable to continue to use a frame that is known to be biased, we decided to adopt PAF as the sampling frame for future British Social Attitudes surveys. For details of the PAF/ER 'splicing' experiment, see Lynn and Taylor (1995).

2. This includes households not containing any adults aged 18 or over, vacant dwelling units, derelict dwelling units, non-resident addresses and other deadwood.

3. In 1993 it was decided to mount a split-sample experiment designed to test the applicability of Computer-Assisted Personal Interviewing (CAPI) to the British Social Attitudes survey series. As the name implies, CAPI involves the use of a laptop computer during the interview, with the interviewer entering responses directly into the computer. There was, however, concern that a different interviewing technique might alter the distribution of responses and so affect the year-on-year consistency of British Social Attitudes data. Following the experiment, it was decided to change over to CAPI completely in 1994 (the self-completion questionnaire still being administered in the conventional way). The results of the experiment are discussed in The 11th Report (Lynn and Purdon, 1994).

4. Interview times recorded as less than 20 minutes were excluded, as these timings were likely to be errors.

5. An experiment was conducted on the 1991 British Social Attitudes survey (Jowell et al., 1992) which showed that sending advance letters to sampled addresses before fieldwork begins has very little impact on response rates. However, interviewers do find that an advance letter helps them to introduce the survey on the doorstep, and a majority of respondents have said that they preferred some advance notice. For these reasons, advance letters have been used on the British Social Attitudes surveys since 1991.

6. Because of methodological experiments on scale development, the exact items detailed in this section have not been asked on all versions of the questionnaire each year.

7. In 1994 only, this item was replaced by: Ordinary people get their fair share of the nation's wealth. *[Wealth1]*

8. In constructing the scale, a decision had to be taken on how to treat missing values ("Don't know" and "Not answered"). Respondents who had more than two missing values on the left–right scale and more than three missing values on the libertarian–authoritarian and welfarism scales were excluded from that scale. For respondents with fewer missing values, "Don't know" was recoded to the mid-point of the scale and "Not answered" was recoded to the scale mean for that respondent on their valid items.

9. See www.bsa-30.natcen.ac.uk.

References

DeVellis, R. (2003), *Scale Development: Theory and Applications*, 2nd edition, Applied Social Research Methods Series, 26, Thousand Oaks, California: Sage

Jowell, R., Brook, L., Prior, G. and Taylor, B. (1992), *British Social Attitudes: the 9th Report*, Aldershot: Dartmouth

Lynn, P. and Purdon, S. (1994), 'Time series and lap-tops: the change to computer-assisted interviewing', in Jowell, R., Curtice, J., Brook, L. and Ahrendt, D. (eds.), *British Social Attitudes: the 11th Report*, Aldershot: Dartmouth

Lynn, P. and Taylor, B. (1995), 'On the bias and variance of samples of individuals: a comparison of the Electoral Registers and Postcode Address File as sampling frames', *The Statistician*, **44**: 173–194

Spector, P. (1992), *Summated Rating Scale Construction: An Introduction*, *Quantitative Applications in the Social Sciences*, 82, Newbury Park, California: Sage

Stafford, R. and Thomson, K. (2006), *British Social Attitudes and Young People's Social Attitudes surveys 2003: Technical Report*, London: NatCen Social Research

At **NatCen Social Research** we believe that social research has the power to make life better. By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to focus our time and energy on meeting our clients' needs and delivering social research that works for society.

Publication details

Park, A., Bryson, C., Clery, E., Curtice, J. and Phillips, M. (eds.) (2013), *British Social Attitudes: the 30th Report*, London: NatCen Social Research, available online at: www.bsa-30.natcen.ac.uk

First published 2013

NatCen Social Research
35 Northampton Square
London
EC1V 0AX

info@natcen.ac.uk

ISBN 978-1-907236-28-0