Centre for Longitudinal Studies ▫ Institute of Education ▫ University of London ▫ London WC1H 0AL

1970 British Cohort Study (BCS70)

Twenty one-year Sample Survey

A Guide to the BCS70 21-year Data

available at the

Economic and Social Research Council Data Archive

***Deposited by*:**

John Bynner,
Centre for Longitudinal Studies
Institute of Education
University of London

Document prepared by:

Sofia Despotidou
Peter Shepherd
Brian Dodgeon

CONTENTS

continued...

Acknowledgements

The preparation for the 1970 Birth Cohort Study (BCS70) was originally laid down in the British Births Study. This was organised and funded mainly by the National Birthday Trust and the Royal College of Obstetricians. The Medical Research Council (MRC) funded sample surveys at 22 and 42 months both directed by Dr Chamberlain. Neville Butler directed the 5-year, 7-year, 10-year and subsequently the 16-year follow-ups. John Bynner has directed subsequent surveys at 21 years (a 10 per cent sample), 26 and 30 years.

The Twenty-one-year Sample Survey was funded mainly by three sources, the Adult Literacy and Basic Skills Unit (ALBSU), The Paul Hamlyn Foundation, and the Leverhulme Trust. Other funding sources include the Dulverton Trust, the International Centre for Child Studies, the Mirror Group Newspapers Plc and the Trustee Savings Bank Plc.

The following SSRU staff were involved in the design of the survey instrument, and the preliminary analysis of the data: John Bynner, Neville Butler, Scott Montgomery, Carol Ekinsmyth and Peter Shepherd. Kevin Dodwell and a number of temporary clerical staff undertook the initial tracing of the cohort members, and the coding of open-ended questions, Scott Montgomery and Carol Ekinsmyth were responsible for the creation of the SPSS dataset and initial data cleaning at SSRU; Sofia Despotidou and Brian Dodgeon undertook further checking and updating.

The MORI research organisation assisted with the development of the survey instrumentation, undertook the pilot and main surveys, coding and initial data entry and data cleaning.

Preface

This document is a guide to the 1970 British Cohort Study (BCS70) Twenty one-year Sample Survey. It is designed to accompany the first deposit of data from this survey with the Economic and Social Research Council Data Archive, at the University of Essex. The documentation and data have been prepared initially by the Social Statistics Research Unit (SSRU) at City University, which took over responsibility for BCS70 in 1991, and subsequently by the Centre for Longitudinal Studies (CLS) at the Institute of Education.

In Section 1, a short history is given of the whole longitudinal study. This is followed by a more detailed account of the Twenty one-year survey begun in 1991. Section 2 provides an in-depth guide to using the 21-year data, outlining amongst other issues, the data cleaning policy, response bias and reliability of specific questions. In Section 3 the annotated questionnaire is reproduced. In Section 4, there are 3 appendices. The first gives a summary of topics covered by BCS70 surveys to date. The second appendix is the list of publications. The third provides the syntax used to generate the derived variables included on the deposited dataset.

Everyone is strongly advised to read Section 1 first, as the historical context facilitates a clearer interpretation of the data.

Summary of the Data Deposit

1970 Birth Cohort Study: Twenty one-year Sample Survey

Principal Investigator	John Bynner, Director, Social Statistics Research Unit, City University, EC1V OHB.
Original Organisation	City University, Social Statistics Research Unit
Depositor	John Bynner, Director, Centre for Longitudinal Studies (CLS), Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL.
Documentation	Prepared at CLS by Sofia Despotidou, Brian Dodgeon and Peter Shepherd.
Acknowledgements	Fieldwork, data entry and initial data cleaning: MORI.
Purpose	To investigate the processes and outcomes of transition from secondary education to employment; and to investigate levels of literacy and numeracy.
Variables	Three survey instruments were used: Self-Completion Questionnaires (Cohort), an interview and literacy and numeracy assessments
Population	Children born 5-11 April 1970.
Sampling Procedures	The sample was drawn from cohort members who where resident in England and Wales. Interviews conducted in 25 clusters based on 26 post code areas.
Cases (target)	More than 10,000 (those with current address available)
Cases (obtained)	1,647
Method of Data Collection	Self-completion questionnaires, a face-to-face interview and an assessment
Date of Data Collection	1992

SECTION 1

INTRODUCTION TO THE 1970 BRITISH COHORT STUDY, WITH PARTICULAR REFERENCE TO THE TWENTY ONE-YEAR SAMPLE SURVEY

INTRODUCTION

1.1 Britain has a unique tradition in conducting longitudinal birth cohort studies. Four continuing studies have been embarked upon. These were in 1946 (National Survey of Health and Development, NSHD), in 1958 (National Child Development Study, NCDS), in 1970 (1970 British Cohort Study, BCS70), and in 2001 (Millennium Cohort Study). The first three were launched as a study of antenatal/postnatal service provision, and morbidity or perinatal mortality. All those three studies collected information about almost all births occurring nationwide in a target week in 1946 (NSHD, n=13,687), in 1958 (NCDS, n=17,414), and in 1970 (BCS70, n=17,198) respectively. Each of those studies has subsequently carried out multidisciplinary follow-up surveys (or sweeps) on health, education, family and social influences at various ages (Douglas, 1964; Douglas et al 1968; Wadsworth 1979, 1991; Butler and Bonham, 1963; Butler and Alberman, 1969; Davie et al, 1972; Fogelman, 1976, 1983; Ferri, 1993; Chamberlain et al, 1973, 1975; Osborn et al, 1984). The studies present individually and in combination an unprecedented opportunity to investigate the forces and patterns that have shaped and continue to shape the lives of three overlapping generations of people living in Great Britain. The Millennium Study will add a longitudinal dimension to a fourth cohort in the coming years, when follow-up sweeps have been completed.

1.2 This document is designed to accompany the deposit of data from the ***BCS70 Twenty one-year Sample Survey*** with the Economic and Social Research Council Data Archive at the University of Essex. The deposit of this data means that the Data Archive now holds data for the complete longitudinal record from birth to thirty-year data from the other BCS70 follow-ups having been deposited some time ago.

1.3 The present section is in two parts. The first provides details of the background to BCS70 in general, and describes the availability of the data, and some of the publications arising from research using the BCS70 data. The second, focuses on the BCS70 Twenty-one-year Sample Survey in particular, providing details of the design and conduct of the survey, and the initial data preparation. Later sections provide a guide to using the 21-year data, and an annotated copy of the survey instrument employed. An outline of additional data cleaning and documentation undertaken by the Centre for Longitudinal Studies (CLS) is also included.

The documentation and data have been prepared by the Centre for Longitudinal Studies at the Institute of Education, which is responsible for BCS70, as well as for NCDS and the Millennium Cohort Study.

BACKGROUND TO THE 1970 BRITISH COHORT STUDY

Outline of the Longitudinal Study

1.4 BCS70 began in 1970 when data were collected about the 17,198 babies born in England, Scotland, Wales and Northern Ireland in the week 5-11 April. At this time, the study was named the British Births Survey (BBS) and it was sponsored by the National Birthday Trust Fund, in association with the Royal College of Obstetricians and Gynaecologists. Since 1970, there have been five attempts to gather information from the full cohort - when they were aged 5, 10, 16, 26 and 30 years.

1.5 As Table 1.1 shows, information has been obtained from a number of different sources, and by varying types of instruments. With each successive attempt, the scope of enquiry has broadened from a strictly medical focus at birth, to encompass physical and educational development at the age of five, and physical educational and social development at the age of ten. This wider and more comprehensive approach has been continued in the 16-year, 26-year and 30-year sweeps.

1.6 In the birth survey, information was collected by means of a questionnaire completed by the midwife present at the birth, and supplementary information was obtained from clinical records. The five-year and ten-year follow-ups were carried out by the Department of Child Health, Bristol University and the survey at these times was named the Child Health and Education Study (CHES). On both occasions parents of the cohort members were interviewed by Health Visitors, the children themselves undertook tests of ability, and the school health service gathered medical information on each child. This was supplemented at ten years by information gathered on a questionnaire completed by head and class teachers, and the children had

completed a medical examination. In 1975 and 1980, the cohort was augmented by the inclusion of immigrants to Britain who were born in the target week in 1970. Subjects from Northern Ireland who had been included in the birth survey, were dropped from the study in all subsequent sweeps.

The sixteen-year follow-up was carried out by the International Centre For Child Studies and named Youthscan. In this sweep, sixteen separate survey instruments were employed, including parental questionnaires, head and class teacher questionnaires and medical examinations. The cohort members completed questionnaires, kept two four-day diaries and undertook some educational assessments.

The twenty-six year follow-up was carried out by the Social Statistics Research Unit, City University and it was now called the BCS70 26-year survey.

The 30-year survey was carried out by the Centre for Longitudinal Studies after the SSRU moved to the Institute of Education.

Table 1.1: BCS70 Follow-ups, sample surveys and sources of information 1970-2000

pBBS (1970) ***Birth***	**BBS*** (1972)** ***22* mths**	**BBS*** (1973)** ***42* mths**	**CHES (1975)** ***5***	**CHES (1980)** ***10***	**Youthscan (1986)** ***16***	**BCS70 (1991)** ***21******	**BCS70 (1996)** ***26***	**BCS70 (2000)** ***30***
Mother	Mother	Mother	Parents	Parents	Parents			
				School	School			
	Test	Test	Test	Tests	Tests	Tests		
Medical	Medical	Medical	Medical	Medical	Medical			
				Subject	Subject	Subject	Subject	Subject
16,135 *	2,457	2,315	13,135	14,875	11,628	1,647	9,003	11,261
98%**	15%	14%	80%	93%	72%	10%	56%	70%

Notes: * Achieved Sample - at least one survey instrument partially completed. (The figure for the BBS comprises cohort members living in GB known to be alive at the end of one week).

** Per cent response.

*** Sample surveys: 22 months, 42 months, and 21 years.

1.7 In the 1972 and 1973 sample surveys the British Births Child Survey took as its subjects three sub-samples: all twins in the original cohort were included, as were low birthweight and post-mature births. A random ten per cent sample of the original cohort acted as a control group. The South-West Region Survey, carried out at the age of $3^{1/2}$ surveyed ninety five per cent of the cohort members who lived in the south west of England and Glamorgan, South Wales. These smaller-scale surveys were undertaken in order to bridge what was a large gap in child development terms between birth and five years of age (when the next full sweep was planned). In 1977, an attempt was made to obtain details about those who had not responded in the five-year survey. A postal questionnaire was completed by parents of 1,917 of the non-responders.

Availability of Data

1.8 Data for the *birth survey,* the *22-month* and *42-month sample surveys,* the *five-year,* the *ten-year, sixteen-year, twenty six-year and thirty-year* follow-ups are already lodged at the *ESRC Data Archive*, and are available to the research community for analysis.

1.9 Data for the *twenty-one year* sample survey accompany this document.

A fully documented database containing all BCS70 data, is now being prepared by Centre for Longitudinal Studies and will be made available via the ESRC Data Archive. Until data are generally available in this way, it is possible to obtain data not already lodged with the Data Archive from the Centre for Longitudinal Studies directly - see *paragraph 1.35* below for more details.

Some Examples of Published Material

1.10 Wide-ranging use has already been made of the data arising from the birth survey and the five follow-ups of BCS70. It is not within the scope of this document to review the literature, but it is useful when considering the background to the project to mention some key works. The birth sweep provided a valuable insight into the patterns at that time of obstetric and neonatal care in the United Kingdom (Chamberlain et al 1973, 1975). The birth and five-year findings relating to health were discussed in Butler, Golding and Howlett's (1986) *From Birth to Five: A Study of the Health and Behaviour of Britain's 5-Year Olds*, and general findings from the first two sweeps were also outlined in a book written by Osborn, Butler and Morris (1984). Chamberlain and Simpson (1979) also concentrated on the health data arising from the birth, 22 and 42 month surveys in their book *The Prevalence of Illness in Childhood*. An account of early findings from the 26-year follow-up is given in *Twenty-Something in the 1990s,* Bynner, et al. (1997).

1.11 In addition to these general accounts of the findings of BCS70, a number of specific projects have been undertaken in a wide range of fields. In the area of health, investigations have ranged from vision problems (Atkinson and Butler, 1985; Stewart-Brown, 1986) to childhood accidents (Bijur, 1984; Wadsworth et al, 1983a), appendicitis (Barker et al, 1988), breast-feeding (Taylor et al, 1983a, 1984) teenage mothering (Taylor et al, 1983b; Wadsworth et al, 1983b) and much more. Special educational problems and needs have formed the scope of some of the research arising from the BCS70 data. For example, Haslum and Butler (1985) considered the special education needs of ten year olds, and Rodgers (1983) investigated the prevalence of reading retardation. Different forms of social behaviour and their influence on health and development have also been investigated. Rush and Cassano (1983) considered the influence of parental smoking on perinatal mortality; Haslum, Morris and Golding (1984) reported on the diets of Britain's ten year olds; Osborn (1984) considered maternal employment and depression and their influence on child behaviour; and Osborn and Morris (1982) investigated fathers' roles in child care. Osborn and Milbank (1987) showed the longer term value of pre-school education and day-care. A full bibliography of publications and reports arising from BCS70 is reproduced in Appendix 3.

1.12 The data already collected provide a remarkably rich research resource in a large number of areas. Future data collection by means of regular sweeps, and the adoption of an integrated approach to the design and analysis of this study with the 1958 cohort study (NCDS) will present exciting new possibilities and dimensions for research.

THE BCS70 TWENTY-ONE-YEAR SAMPLE SURVEY

Rationale and Overall Design

1.13 The ***BCS70 Twenty one-year Sample Survey*** is the fourth sample survey of the 1970 cohort born in Great Britain 5-11 April 1970. The cohort has been surveyed comprehensively at birth, five, ten, 16, 26 and 30 years, and samples were seen at 22 months, 42 months, seven and 21 years. The study was originally titled the *British Births Survey*, then the *Child Health and Education Study* (CHES), then *Youthscan,* then in 1991 renamed the British Cohort Study 1970 (BCS70).

1.14 A number of considerations influenced the development of the survey instrumentation used for the BCS70 Twenty one-year Sample Survey.

They can be said to fall under two broad headings:-

- **Literacy and Numeracy (Basic Skills)**

- **Education, Training and Employment**

1.15 Adult ***literacy and numeracy*** problems are a central concern in modern societies. They affect not only employment prospects and the ability to perform well at work, but how adults function in every area of their social and domestic lives and as citizens. Moreover, they point to serious failings of an education system which, over the years of compulsory schooling, has failed to impart the most basic of all educational skills to a proportion of children. During the decade prior to this sample survey, programmes directed at rectifying reading and numeracy deficiencies in adults were carried out in which the Adult Literacy and Basic Skills Unit (ALBSU) had the main coordinating role. But the problem persisted, not least because schools had yet to achieve the goal of literacy and numeracy for all the children passing through them. Nor had training schemes of the kind that burgeoned in the late 70s and the 80s bridged the gap. Gaining a better understanding of the types of problems adults have and how these develop are essential to making literacy and numeracy teaching effective.

Attitudes towards literacy and its definition have changed over the past century. No longer is it sufficient or desirable to classify people according to a simple dichotomy literate/illiterate on the basis of their ability to accomplish a simple task (like writing their name). Recent studies in this country and abroad have suggested that large percentages of the adult populations of developed countries are unable to apply literacy skills which are advanced enough to meet the varying demands placed upon them in their every day lives. This has serious implications both for the individuals themselves (lifechance) and for society in general (in terms of human resource management).

"Literacy" should be seen as a dynamic and context specific concept. Changing environments place ever-changing demands on the individuals within them. Literacy levels necessary for one person or group of people may be very different to those necessary for another. It is thus important to attempt to develop measures of literacy, which are relevant to the people who will be assessed by them. Obviously, the more generally applicable the tests are made, the less personal they can be. It is important therefore also to collect information on people's own views of their abilities and their shortcomings.

The first objective in this sphere was to obtain an estimate of the proportion of English and Welsh young adults who had experienced any reading, writing, spelling or numeracy problems in their adult lives. This was, in part, directed at updating information collected eleven years previously from 23-year olds in the National Child Development Study - the cohort study based on the sample of approximately 17,000 people born in a single week in 1958 (ALBSU, 1987).

Thirteen percent of the NCDS cohort members said that they had had literacy or numeracy problems since leaving school. This BCS70 sample survey offered the opportunity to repeat the questions in order to obtain fresh estimates and see whether there was any evidence of a marked change between the two surveys (such comparisons could, of course, only be crude, because of the differences in age and in the constitution of the sample between the two studies).

The second objective was to develop an assessment scheme for adult reading, writing and numberwork difficulties, which could be administered to respondents by professional survey interviewers. This would enable comparisons to be made between BCS70 cohort members' subjective appraisals of their difficulties and an objective assessment based on performance in a number of everyday tasks. The practical limitations on the design were that the assessment should take no more than 30 minutes to complete, should be simple to administer and be as enjoyable and unthreatening as possible to interviewees. The assessment needed to cover a range of performance at the four levels for communication skills and the three levels for numeracy as defined by ALBSU's Basic Skills Standards. These emphasise 'functional' performance, i.e. the ability to apply basic skills in every day life situations (ACACE, 1982).

The third objective was to investigate the kinds of difficulty people *claimed* to have and were *found* to have from the assessments. Are there particular types of task that people find exceptionally difficult and do these differ between men and women? Time limitations would of course restrict this part of the study to a very limited range of examples.

The fourth objective was to investigate the present and past circumstances and other personal characteristics of people found to have difficulties. This would bring in data from other parts of the survey and from the earlier stages of the study, exploiting its longitudinal features.

1.16 In looking at ***Education, Training and Employment*** in the context of a survey of 21-year-olds, it is important to note how routes to employment for young people have evolved in recent decades. Apprenticeship was traditionally viewed as the main way that young people learnt the skills and amassed the experience necessary to enter an occupation. As apprenticeships became rarer, other forms of education and training became more important in the transition from school to work. Although education and training at schools and colleges is invaluable, acquiring work experience and learning 'on the job' is increasingly important.

With increasing age, an individual's ability to influence their career path may diminish. Someone with poor school qualifications may find it difficult to secure a desirable job or to get the necessary training and experience. Relevant courses may not be available locally, or there might be financial or other personal constraints that prevent young people enhancing their qualifications and abilities. Occupational aspirations of some young people may not be matched by their examination performance at school. What are the vital factors that contribute to relative failure? Undeniably, those who leave school with few qualifications will have fewer courses open to them than those who are well qualified. To what extent does poor performance at school persist in influencing the adult working life? In investigating the routes through postcompulsory education and training it is possible to identify the strategies or circumstances that lead to a relatively successful career route, despite early educational failure.

(a) Youth Training: In developing a typology, Youth Training was of particular interest because it should represent a genuine transition stage between full-time education and employment. YT has been modified since it was introduced by the government with the objective of improving young people's employment prospects (Department of Education and Science, 1981; Department of Employment, 1986) and providing them with the skills required for modern industry (Department of Employment, 1988). YTS began as a one year scheme designed to reduce youth unemployment and was extended to two years in 1978. The 1988 Employment Act replaced the national scheme with Youth training, (YT) which was to be run by the newly established Training and Enterprise Councils (TECs). BCS70 cohort members experienced both one year and two year YT/YTS schemes.

The variation in type of YT available to cohort members requires some consideration. Differences in local labour market skill requirements should be taken into account (Ashton and Maguire, 1986). These differences in skill requirements should affect the kind of training that is offered. Training and Enterprise Councils were set up by the government so that local training requirements could be met (Department of Employment, 1988). This diversity in the provision of YT has made it undesirable to classify everybody who experienced YT in a single YT category.

Three YT sectors were identified by Roberts and Parsell (1989): the sponsored sector, the contest sector and sink schemes, which can be defined as follows.

Trainees who entered the sponsored sector were the most likely to find employment at the end of the scheme through qualifications or apprenticeships. Possibly more importantly, trainees in this sector had more chance of a permanent job with the sponsoring employer when the scheme was over. The contest sector schemes could also provide qualifications and work experience, although there was less certainty of achieving permanent employment with the firm that provided the training. The third YT sector, described as sink schemes by Roberts and Parsell (1989), tended to be community-based rather than employer-led: there were virtually no opportunities for trainees on these programmes to be offered employment when their schemes came to an end.

The different YT sectors tended to provide differing levels of training quality: highest quality training being provided by the sponsored and contest sectors, while the sink schemes offered the least to their trainees. These differences were reflected in recruitment, as those with the poorest qualifications tended to enter the sink schemes. The differences between the sectors were also reflected in the training for qualifications that was provided for trainees. Sink schemes were less likely to provide or lead to qualifications for their trainees.

(b) Full-Time Education: Recruitment into the different categories of the typology is determined by a combination of choice and opportunity (Roberts and Parsell, 1988). It has been argued that although the ambitions of school leavers can shape their careers, the consequences of their decisions were trivial in comparison to other factors such as their educational attainment and the jobs that are available in the local labour market. Roberts and Parsell also supported the argument that further constraints were made on opportunities when some choices were made. For instance, a decision to leave fulltime education may be practically irrevocable after an individual has entered the labour market. It is therefore important to describe the categories of the typology in terms of school qualifications and age of leaving full-time education.

(c) Post-School Vocational Qualifications: The most commonly achieved post-school vocational qualifications included were: Royal Society of Arts awards (RSA); City and Guilds; Joint Industry Board and National Joint Council awards (JIMNC); National Diplomas and Certificates; TEC BEC and BTEC awards (TEC/BEC). Other types of qualification are excluded due to the relatively small numbers.

(d) Unemployment: Youth unemployment is one of the most important outcomes in measuring failure in transition from education to employment. Unemployment is defined here as not having a job and seeking work. Only those who have had one or more periods of three consecutive months of unemployment were categorised as having experienced a significant period of unemployment.

(e) Current or Most Recent Employment: Respondents were asked about the main thing, other than the minimum entry qualifications for the job, that helped them obtain that position. The available response categories for this question were: experience, contacts, other qualifications, specific skills or abilities, or 'something else'.

Tracing

1.17 The continued success of longitudinal studies like BCS70, is dependent on knowing the whereabouts of the study subjects. For the purposes of the follow-ups in 1975, 1980 and 1986, the cohort members were traced mainly through schools, and no effort was made to maintain contact in the intervening years. Following the 1986 survey, it was no longer possible to trace the 16,000 cohort members anew in this way, and after responsibility for the study was assumed by SSRU in 1991, efforts were made to maintain contact with those for whom a current address was still available, by mailing an annual birthday card. This was designed to give cohort members feedback about the use of information they had already given, and to advise them of plans for future surveys. It also provided an opportunity to confirm address details and other helpful information. As a result of these efforts, information about the current whereabouts of some 10,000 of the 16,000 BCS70 cohort members was available at the time that preparations for the 1991 Sample Survey began in autumn 1991.

Survey Instruments

1.18 The *BCS70 Twenty one-year Sample Survey* made use of four questionnaires, "Your Life Since 1986", "Your Views", "Interview", and "Literacy and Numeracy Assessments". This was developed by the cohort studies research team at SSRU, in consultation with those who had been involved with the design and analysis of earlier BCS70 sweeps, and those familiar with the 1958 National Child Development Study (NCDS). It was based on survey instruments used for the 23-year NCDS follow-ups, and was designed to provide details of the views and current circumstances of the cohort members in a number of key areas, plus more limited information on their experiences since the last follow-up in 1986. Successive drafts of the questionnaire were piloted on 50 cohort members which were not part of the sample. The final version included questions concerning: employment and education histories since age 16, qualifications, training,

unemployment, reading and writing behaviour, literacy and numeracy self-appraisal and assessments, household composition, relationships, children, housing, income, health, attitudes to employment, education, literacy and numeracy, and self efficacy.

Sample Design

1.19 The main survey achieved 1,650 interviews using a clustered sample of 25 sampling points based on 26 postcode areas throughout England and Wales. The sample was obtained using interval sampling of all postcode areas in England and Wales with a random starting point. To ensure that the sample would be representative, the characteristics of cohort members resident in the sample areas were compared with those of 21 year olds who participated in the 1989 Labour Force Survey (LFS89). The regional distribution and characteristics of cohort members resident in the sample areas was very similar to the distribution and characteristics of LFS89 respondents.

Pilot

1.20 The instruments and fieldwork procedures were piloted by MORI interviewers between 25th November and 9th December, 1991. The following instruments were piloted: the interview schedule, the self-completion questionnaire and the literacy and numeracy assessments.

Cohort members were initially contacted by telephone (or by personal call where no telephone number was available), an interview date was arranged and a self-completion questionnaire, was then posted to respondents. The questionnaires were collected at the time of interview. If the self-completion questionnaire had not been filled in, interviewers waited for the questionnaire to be completed before starting the interview, as the employment and education history section was required to provide dates. Cohort members were generally very happy to participate in the pilot and were not put off by the initial telephone call or by being asked to complete the questionnaire before the interview.

A total of 85 cohort members' addresses were issued; 33 in South East London and 52 in the Reading area. 49.4% of respondents completed the interview, assessment and self completion questionnaire; 4.7% refused to participate in the survey and 45.9% could not be contacted or were unavailable for the duration of the pilot. The proportion of cohort members who completed the interview was lower than might be expected. This was partly due to the short duration of the fieldwork period, which made it difficult to arrange a convenient time to interview some cohort members or to contact those cohort members who had moved house or were away at college/university. The majority of those who were contacted were very willing to participate.

The mean completion time for the literacy and numeracy assessments was 28 minutes. The average duration of the interview was 54 minutes. Although the assessments were accomplished within the designated time (30 minutes), the interview was, on average, nine minutes longer than planned. A small number of cohort members could not complete the self-completion questionnaire due to literacy difficulties and were assisted by the interviewer.

The pilot was successful, indicating the suitability of fieldwork procedures and questionnaire design, which required only minor modifications for use in the main survey. Initial concern about the literacy and numeracy assessments was unjustified as both interviewers and respondents appeared to be comfortable with the assessments. Overall, The numeracy tests presented more problems to respondents than the literacy assessments.

Data Collection

1.21 Fieldwork for the survey was successfully conducted by MORI between February and May, 1992. MORI were also responsible for data entry. All data collected by the survey was delivered to SSRU on magnetic media.

All initial contact with cohort members made by interviewers was by telephone or personal visit to their home. This procedure was adopted to minimise non-response amongst cohort members, especially those with literacy difficulties. Interviewers were instructed to persuade cohort members to agree to an Interview, arrange a date and time when this would be carried out, and then explain that the two self-completion questionnaires would be sent to them by post. Although the two questionnaires were designed to be completed before the Interview was conducted, it was stressed to cohort members that they did not have to finish completing them before the interviewer arrived. Where cohort members experienced difficulty in answering the questions in the 'Your Life Since 1986' questionnaire, the interviewer completed, or assisted in the completion, of this instrument before the Interview was conducted. The 'Your Life Since 1986' questionnaire was used as a reference document by interviewers throughout the interview to ensure that dates and reported activities were consistent.

A postage-paid envelope was left with those respondents who had not completed the 'Your Views' questionnaire, so that it could be returned at a later time. Interviewers were instructed to explain the meanings of questions in the Your Views questionnaire, but not to help complete it, so that interviewer would not influence the cohort members' answers to attitude questions.

The Literacy and Numeracy Assessments were administered after the Interview was completed.

Fieldwork outcome

1.22 The target of 1,650 interviews was achieved, representing 70 per cent of the 2,359 addresses which were included in the sample. A breakdown of response is given below.

Table 1.2: Interview response

Outcome	%	Number
Interview completed	70	1,650
Refusal	6	141
Respondent had moved to different area	6	140
Respondent had moved – no new address	7	158
No contact	7	176
Incapable of doing interview	-	5
Other	4	89
TOTAL	100	2,359

The largest cause of non-response was that cohort members had moved house in the 9 months prior to fieldwork. Due to the short duration of the fieldwork period and the clustered nature of the sample, it was not possible to pursue interviews with the majority of cohort members who had moved house. The low refusal rate of 6 per cent indicates the extremely cooperative nature of BCS70 cohort members.

Details of the distribution of the sample by region and postcode area are given in Table 1.3 below.

Table 1.3: Distribution of the BCS70 sample by region and postcode area

Region	Postcode area	Planned sample %	Interviews achieved %
North	NE	6.31	6.24
North West	LA	2.92	2.18
	WA	5.24	5.58
Mersey	L	3.22	2.18
Manchester	M	3.03	3.52
West Yorkshire	WF	4.35	4.55
Yorkshire & Humberside	YO	3.37	3.88
South Yorkshire	S	3.17	3.33
East Midlands	NG	4.12	4.30
	NN	2.54	2.73
Anglia	NR	5.03	5.45
South East	SS	3.98	4.61
	RH	3.59	3.45
	SO	4.53	4.67
	OX	4.83	4.79
	DA	2.69	2.79
London	UB	2.84	3.09
	E/SE	6.24	5.94
	W	2.25	1.88
South West	BH	2.56	2.97
	BS	5.50	5.15
Wales	SA	6.42	5.27
West Midlands	ST	2.53	2.61
W Midlands Conurbation	B	6.57	6.91
	WV	1.77	1.94
Total %		100.00	100.00
Total n		1,650	1,650

Data Processing

1.23 Completed questionnaires were returned by cohort members to MORI, who were responsible for a preliminary visual edit agreed with SSRU. Following data entry of all pre-coded information, a preliminary computer edit (again agreed with SSRU) was carried out by MORI to check that data was valid (i.e. single-coded, 0-9), and within range (i.e. as specified in the questionnaire).

1.24 Qualifications were coded using a scheme based on that used for recent NCDS follow-ups, but because of an improved approach to the survey also provided information on each qualification identified.

Sample representativeness

1.25 The 1989 Labour Force Survey (LFS) has been used as a basis for comparison as it claims to be representative of the population in England and Wales, and because we were able to select information for a narrow age band of respondents (21-year-olds). The employment characteristics and marital status of respondents in the BCS70 21-year survey and the LFS are compared below.

The tables show that, on the basis of economic activity and marital status, the two samples are broadly similar. The slight difference between the LFS and BCS70-21 economic activity data can be largely explained by differences in coding conventions. A slightly higher proportion were cohabiting in the BCS70 sample and females outnumber males by 53 to 47 per cent, reflecting expected patterns of response bias in BCS70 data. On the basis of comparison with the LFS data, we are confident that the sample can be used to conduct reliable research on 21 year olds living in England and Wales.

Table 1.4 Economic activity of respondents

Economic activity	LFS %	BCS70-21 %
Employed	67	70
Training scheme	1	1
Looking for work	7	9
Full-time education	9	10
Keeping house	7	7
Other	9	2
TOTAL &	100	100
TOTAL n	1,869	1,650

Table 1.5 Economic activity of respondents

Marital status	LFS %	BCS70-21 %
Married	15	13
Cohabiting	8	13
Single	77	73
Separated/divorced	1	-
TOTAL %	100	100
TOTAL n	1,869	1,650

Response bias

1.26 Although a generally acceptable response rate has been achieved, anything less than a perfect response raises the question of whether those who completed a questionnaire are representative of the sampled population - in this case, the cohort members living in Great Britain (England, Wales and Scotland) for whom SSRU had a current address. This issue has been explored by exploiting a possibility only available to longitudinal studies: comparison of the achieved sample - those cohort members who returned a questionnaire - with the target sample - the cohort at birth, and in subsequent follow-ups.

1.27 Comparisons between the achieved sample and the target sample have been extensive. They are based on variables selected from the earlier BCS70 follow-ups. The variables chosen include many relating to

demography, education, literacy and numeracy problems, social and economic circumstances, financial problems, the family and relationships, housing and household, and health.

1.28 Table 1.7 reports on just some of the comparisons made. It contrasts the characteristics of the target sample and those of the achieved sample for the 21-year follow-up, using a range of variables selected to represent the areas identified above. The absolute difference between the target and achieved samples, and the percentage bias are reported for each variable, indicating the extent of the difference between the cohort and the sample. For this table percentage bias is calculated as follows:

((Per cent in the achieved sample - Per cent in the target population)/Per cent in the target population) x 100

A negative percentage bias indicates under-representation in the 21-year follow-up, and a positive percentage bias shows over-representation. When looking at the contrasts for variables taken from the 1975, 1980, and 1986 follow-ups it is important to remember that the target percentage will itself reflect differential response to these surveys.

1.29 Nevertheless, the analysis provides an important, and generally encouraging, insight into differential response. Absolute differences between the sampled population and the achieved sample are, on the whole, small and this is reflected in many of the figures for percentage bias. However, small absolute differences can result in a relatively large figure for percentage bias where the percentage in the sampled population is small. Levels of statistical significance are not reported, but it should be noted that, in samples of this size, tests of statistical significance are sensitive to very small differences. In general, the achieved sample does not differ greatly from the sampled population.

1.30 Overall, it appears that men, those born outside Britain, and those with minority ethnic background are under-represented in the BCS70 21-year follow-up. This also holds for those born to single mothers, teenage mothers, and unemployed fathers; as well as for those with low school achievement; those with a disability, those who have been in care, those with origins in the lower social classes, those who grew-up in families with financial problems, and those who have experienced poor housing conditions.

1.31 Perhaps not surprisingly, this picture is similar to that emerging from the analyses of differential response to other surveys, and especially the NCDS follow-ups.

Table 1.6: Response bias – by Region

Region	***Planned***	***Achieved***	***Difference***	***Bias***
	%	**%**	**%**	**%**
North	6.31	6.24	-0.1	-1.1
North West	8.16	7.76	-0.4	-4.9
Mersey	3.22	2.18	-1.0	-32.3
Manchester	3.03	3.52	0.5	16.2
West Yorkshire	4.35	4.55	0.2	4.6
Yorkshire & Humberside	3.37	3.88	0.5	15.1
South Yorkshire	3.17	3.33	0.2	5.0
East Midlands	6.66	7.03	0.4	5.6
Anglia	5.03	5.45	0.4	8.3
South East	19.62	20.31	0.7	3.5
London	11.33	10.91	-0.4	-3.7
South West	8.06	8.12	0.1	0.7
Wales	6.42	5.27	-1.2	-17.9
West Midlands	2.53	2.61	0.1	3.2
W Midlands Conurbation	8.34	8.85	0.5	6.1

Table 1.7: Response bias – the sample compared with earlier BCS70 follow-ups

	Cohort %	*Sample* %	*Difference* %	*Bias* %
Male respondents	51.8	45.7	-6.1	-11.8
1970 Birth Survey				
Mother born outside Britain	10.5	8.3	-2.2	-21.0
Father born outside Britain	11.1	9.2	-1.9	-17.1
Mother completed education aged less than 15 years	7.0	6.5	-0.5	-7.1
Father completed education aged less than 15 years	8.8	6.8	-2.0	-22.7
Teenage mother	9.7	8.5	-1.2	-12.4
Single mother	5.8	3.8	-2.0	-34.5
Premarital conception	8.2	8.2	0.0	0.0
Twin at birth	2.2	1.8	-0.4	-18.2
Father's social class - Manual	64.7	67.7	3.0	4.6
Father unemployed	3.5	2.2	-1.3	-37.1
1975 Follow-up				
Child's ethnic group - West Indian	1.2	1.4	0.2	16.7
Parents have no educational qualifications	40.0	39.0	-1.0	-2.5
Mother aged over 40 years	2.3	2.2	-0.1	-4.3
Child living with both natural parents	90.2	91.7	1.5	1.7
Mother and child ever separated for one month or more	5.0	4.7	-0.3	-6.0
Father's social class - Manual	64.7	67.8	3.1	4.8
Weighed under 5lbs at birth	6.9	7.1	0.2	2.9
No congenital abnormality	93.0	93.5	0.5	0.5
No disability	93.2	93.4	0.2	0.2
Family has moved 3 or more times since 1970	10.6	8.7	-1.9	-17.9
Accommodation is crowded (> 1 person/room)	17.9	14.5	-3.4	-19.0
Accommodation rented privately	6.3	6.3	0.0	0.0
Social rating of neighbourhood - Poor	8.2	7.5	-0.7	-8.5

continued...

Table 1.7: Response bias – the sample compared with earlier BCS70 follow-ups (continued)

	Cohort %	*Sample %*	*Difference %*	*Bias %*
1980 Follow-up				
Has great difficulty with mathematics	3.8	2.8	-1.0	-26.3
Has great difficulty with reading	3.4	2.2	-1.2	-35.3
Has great difficulty with writing	2.7	2.0	-0.7	-25.9
Child has lived with the same parents since birth	85.5	87.1	1.6	1.9
Child has lived with only one parent	17.7	14.7	-3.0	-16.9
Child living in residential institution	0.7	0.4	-0.3	-42.9
Father's social class - Manual	53.5	54.4	0.9	1.7
Family receives Supplementary Benefit	9.7	6.6	-3.1	-32.0
Has a disability	7.9	7.5	-0.4	-5.1
Child has never smoked	74.0	75.1	1.1	1.5
Family lived at 4+ places since 1975	14.3	13.3	-1.0	-7.0
Accommodation rented privately	3.3	4.0	0.7	21.2
1986 Follow-up				
Child born outside Britain	2.4	2.2	-0.2	-8.3
Pupil's reading ability impaired	9.2	6.7	-2.5	-27.2
Child assessed as top of academic ability range	4.9	3.7	-1.2	-24.5
Child wishes to stay at school to take 'A' levels	44.4	38.5	-5.9	-13.3
Child lives with natural parents	80.5	82.1	1.6	2.0
Child has been "in care"	1.4	1.2	-0.2	-14.3
Child wants to leave home as soon as possible	5.4	3.2	-2.2	-40.7
Father's social class - Manual	49.2	54.6	5.4	11.0
Family receives Supplementary Benefit	12.2	11.8	-0.4	-3.3
Has disability	10.3	6.1	-4.2	-40.8
Child does not smoke	78.3	79.3	1.0	1.3
Four+ addresses since 1980	1.9	1.1	-0.8	-42.1
Accommodation is rented privately	2.8	2.5	-0.3	-10.7

Cohort % = Per cent in BCS70.
Sample % = Per cent in achieved sample.
Difference = Sample %-Cohort %.
Bias % = ((Sample % - Cohort %)/Cohort %) x 100.

Further Information on BCS70

1.32 For more information about the BCS70 Twenty-one-year Sample Survey, contact the BCS70 User Support Group by post, telephone, fax, or email as shown below:

Post: BCS70 User Support Group
Centre for Longitudinal Studies
Institute of Education
University of London
20 Bedford Way
London WC1H 0AL

Tel: 020-7612-6860

Fax: 020-7612-6880

Email: cohort@cls.ioe.ac.uk

User Support Group

1.33 The *BCS70 User Support Group* provides advice and guidance on the use of BCS70 data; produces documentation; collates and disseminates information on uses of the data, publications, and other developments; produces and distributes a newsletter and working papers; provides access to non-computerised BCS70 data; collects additional information; and services the User Group.

User Group

1.34 The *BCS70 User Group* is open to all users of BCS70 data. It provides opportunities for users to get together to explore developments, problems, and other issues of mutual interest. Ad hoc "Updates" on BCS70 data and developments are circulated to members.

Membership is free on application to the User Support Group.

Acquiring BCS70 Data for Research

1.35 As noted above, data sets containing the birth, 22-month, 42-month, 5-year, 10-year, 16-year, 21-year, 26-year and 30-year data are already lodged at the ESRC Data Archive, and are available to the research community for analysis.

1.36 A fully documented longitudinal database, which will contain all BCS70 data, is also being prepared by CLS and, when complete, a copy will also be made available via the ESRC Data Archive. Until data are generally available in this way, it is possible to obtain data not already lodged with the Data Archive from the CLS directly (see *paragraph 1.35* above). In the meantime, longitudinal datasets may be created by researchers by merging data from the individual follow-ups already held in the ESRC Data Archive (see *paragraph 2.13* below).

The ESRC Data Archive may be contacted by post, telephone, fax, or email as shown below:

Post: ESRC Data Archive
University of Essex
Colchester
CO4 3SQ

Tel: (01206) 872001

Fax: (01206) 872003

Email: archive@essex.ac.uk
www.data-archive.ac.uk

References

1.37 Details of the publications cited in this section are given below. A full list of publications arising from BCS70 is given in *Appendix 3*.

Atkinson, S. and Butler, N.R. (1985) Vision Problems in Under 5s In *The At-Risk Infant: Psycho/Social/Medical Aspects*. Harel, S. and Anastasia, N.J. (eds). London: Paul Brookes Publishing Co.

Banks, M., Bates, I., Breakwell, G., Bynner, J., Emler, N., Jamieson, L., and Roberts, K (1992) *Careers and Indentities.* Buckingham: Open University Press.

Barker, D.J.P., Osmond, C., Golding, J. and Wadsworth, M.E.J. (1988) Acute appendicitis and bathrooms in three samples of British children. *British Medical Journal*, 262 pp. 956-958.

Beck, U. (1986) *Risk Society.* London: Sage.

Bijur, P. (1984) *The Relationship Between Child Behaviour and Accidents from Birth to Five: a multivariate analysis of a national cohort.* Ph.D Thesis Columbia University, New York.

Butler, N.R. and Bonham, D.G. (1963) *Perinatal Mortality*. Edinburgh: E & S Livingstone.

Butler N.R., and Alberman, E.D. (1969) *Perinatal Problems*. Edinburgh: E & S Livingstone.

Butler, N.R., Golding, J. and Howlett, B. (1986) *From Birth to Five: A study of the health and behaviour of Britain's 5-year olds*. Oxford: Pergamon Press.

Bynner, J. and Ashford, S. (1994) Politics and participation: Some antecedents of young people's attitudes to the political system and political activity. *European Journal of Social Psychology,* 24, pp223-236

Bynner, J., Ferri, E. and Shepherd, P. (Eds) (1997) *Twenty-something in the 1990s: getting on, getting by, getting nowhere.* Aldershot: Ashgate

Chamberlain, R., Chamberlain, G., Howlett, B.C. and Claireux, A. (1973) *British Births: Vol. 1. The First Week of Life*. London: Heinemann Medical Books.

Chamberlain, R., Chamberlain, G., Howlett, B. and Claireux, A. (1975) *British Births 1970, Vol. 2. The First Week of Life*. London: Heinemann Medical Books.

Chamberlain, R.N. and Simpson, R.N. (1979) *The Prevalence of Illness in Childhood.* Tunbridge Wells: Pitman Medical.

Davie, R., Butler, N., and Goldstein, H. (1972) *From Birth to Seven*. London: Longman.

Douglas, J.W.B. (1964) *The Home and the School*. London: MacGibbon & Kee.

Douglas, J.W.B., Ross, J.M. and Simpson H.R. (1968) *All Our Future: A Longitudinal Study of Secondary Education*. Peter Davies.

Evans, K. and Furlong, A. (1997) Metaphors of youth transition: Niches, pathways, trajectories and navigation. In: Bynner, J., Chisholm, I., and Furlong, A. (eds) *Youth, Citizenship and Social Change.* Aldershot: Ashgate.

Ferri, E. (Ed) (1993) *Life at 33: the fifth follow-up of the National Child Development Study*. London: National Children's Bureau.

Fogelman, K. (1976) *Britain's Sixteen-Year-Olds*. London: National Children's Bureau.

Fogelman, K. (Ed) (1983) *Growing Up in Great Britain: collected papers from the National Child Development Study*. Macmillan.

Galbraith, J. K. (1992) *The Politics of Contentment.* London: Sinclair Stevenson.

Haslum, M.N.; Morris, A.C. and Golding, J. (1984) What do our ten year old children eat? *Health Visitor*, 57, pp. 178-9.

Haslum, M.N. and Butler, N.R. (1985) A cohort study of special educational needs in ten-year olds in the United Kingdom. In *Understanding Learning Disabilities International and Multidisciplinary Views*. Duane, .D. and Leong, C.K. (eds). New York: Plenum Press, pp. 18-26.

Hutton, W. (1995) *The State We're In.* London: Cape.

Osborn, A.F. and Morris, A.C. (1982) Fathers and childcare. *Early Child Development and Care*, No. 8, pp. 279-307.

Osborn, A.F. (1984) Maternal employment, depression and child behaviour. *Equal Opportunities Commission Research Bulletin*, No. 8, pp. 48-67.

Osborn, A.F., Butler, N.R. and Morris, A.C. (1984) *The Social Life of Britain's Five-year-olds: A Report of the Child Health and Education Study.London:* Routledge and Kegan Paul.

Osborn, A.F. and Milbank, J.E. (1987) *The Effects of Early Education. A report from the Child Health and Education Study of children in Britain born 5-11 April 1970.* Oxford: Clarendon Press.

Rodgers, B. (1983) The identification and prevalence of specific reading retardation. *British Journal of Educational Psychology*, 53, pp. 369-73.

Rush, D. and Cassano, P. (1983) Relationship of cigarette smoking and social class to birthweight and perinatal mortality among all births in Britain, 5-11 April 1970. *Journal of Epidemiology and Community Health*, 37, pp. 249-55.

Stewart-Brown, S. (1986) Vision defects in school-children: Screening policy and educational implications. In *Progress in Child Health, Vol. 3.* Macfarlane, J.A. (ed). London: Churchill Livingstone.

Taylor, B., Wadsworth, J., Golding, J. and Butler, N.R. (1983a) Breastfeeding, eczema, asthma and hay fever. *Journal of Epidemiology and Community Health*, 37, pp. 95-9.

Taylor, B., Wadsworth, J., Golding, J. and Butler, N.R. (1983b) Teenage mothering: admission to hospital and accidents during the first five years. *Archives of Disease in Childhood,* 58, pp. 6-11.

Taylor, B., Wadsworth, J., Wadsworth, M.E.J. and Peckham, C.S. (1984) Changes in the reported prevalence of childhood eczema since the 1939-45 war. *The Lancet* (ii) pp. 1255-1257.

Wadsworth, J., Burnell, I., Taylor, B. and Butler, N.R. (1983a) Family type and accidents in pre-school children. *Journal of Epidemiology and Community Health*, 37, pp. 100-4.

Wadsworth, J., Taylor, B., Osborn, A. and Butler, N.R. (1983b) Teenage mothering: child development at five years. *Journal of Child Psychology and Psychiatry*, 25, pp. 305-13.

Wadsworth, M.E.J. (1979) *Roots of Delinquency: Infancy, Adolescence and Crime*. Oxford: Martin Robertson.

Wadsworth, M.E.J. (1991) *The Imprint of Time: Childhood, History and Adult Life*. Oxford: University Press.

Wilkinson, H. (1994) *No Turning Back: Generations and the Genderquake.* London: Demos.

SECTION 2

USING THE 21-YEAR DATA

INTRODUCTION

2.1 As mentioned in Section 1, the ***BCS70 Twenty one-year Sample Survey*** made use of self-completion questionnaires, an interview and assessments. An introduction to the development and use of this document has been outlined in *Section 1*. A copy of the questionnaire is reproduced in Section 3.

2.2 This section is designed to assist users of the data by describing:

- Variable names
- Data cleaning undertaken by CLS
- The coding of missing values
- The derived variables included on the dataset
- An introduction to using the Interactive Data Dictionary
- Linkage of ***BCS70 Twenty one-year Sample Survey*** data to data from other sweeps

VARIABLE NAMES

2.3 The v***ariable names*** for the BCS70 Twenty one-year dataset have been allocated as follows:

- ***Precoded questions*** - variable names are based on card and column numbers taken form the card images used for data entry.
- ***Derived variables*** - variable names are those allocated by those responsible for the code. They are most commonly mnemonics.

Details of all variable names are to be found on the annotated questionnaire in Section 3 below.

CLEANING OF THE BCS70 21-YEAR DATA BY SSRU AND CLS

2.4 As noted above, the MORI research organisation who distributed the questionnaires, were also responsible for a preliminary visual edit of each questionnaire; and following data entry of all pre-coded information, for a preliminary computer edit to check that data was valid (i.e. single-coded, 0-9), and within range (i.e. as specified in the questionnaire). Both visual and preliminary computer edit were agreed with SSRU.

2.5 The work undertaken by SSRU, and later by CLS, has concentrated on the setting-up of an SPSS dataset; and the completion of additional checking and updating. The latter has focused, mainly, on repeating and extending the checks for validity and range, and carrying out new checks for consistency.

A brief outline of the approach to data cleaning is given below.

Overall data cleaning policy

2.6 Extensive checks of data validity, range and consistency were carried out. To be valid the value for any data item for any case must be ***numeric*** (a number, 0-9, etc) or *alphanumeric* (text, A-Z/a-z, etc) as specified in the questionnaire/coding frame/data definition. To be *within range*, the value for any data item for any case must be within the range of values specified in the questionnaire/coding frame/data definition. To be *consistent* the value of any data item for any case must conform to the filter structure of the survey instrument.

2.7 The data for the Twenty one-year Sample Survey have been cleaned through computer editing, using checks for validity, range, and consistency derived from the questionnaire. Due to limited resources and record storage problems, it has often been impossible at this stage to check 'errors' against the original questionnaires themselves. As a result, the 'errors' have been 'corrected' according to the rules outlined in Table 2.1 below. Thus, the identification of 'errors' was followed by a check of the raw data to identify data entry problems. For 'errors' of range and consistency, this was followed by review of the editing rules. Consistency 'errors' were investigated further by back-checking the consistency of the primary filter, and updating this if necessary. Where updates to the edit and/or values had been made, the edit was re-run to check that the 'corrections' had been correctly applied and no new 'errors' introduced. Where no correction could be identified, the 'error' value was set to missing. All edits were iterated until no 'errors' were identified.

2.8 It is important to stress that this procedure differs in one important aspect from that adopted for the cleaning of the BCS70 16-year data which is deposited with the ESRC Data Archive**.** For this dataset *forward coding* was employed to 'correct' filter 'errors'. With *forward coding*, if the primary question is answered in a way that does not lead on to the subsidiary questions then any responses in the subsidiary sections are set to a missing value. For example, if the question is 'Have you ever been to hospital?' then only those that have ticked 'Yes' will be included in the subsidiary questions on 'date', 'reason', etc for admittance to hospital. With an interview based questionnaire, this is obviously what would happen during the interview. However, with self-completion questionnaires, sometimes respondents change their minds when they see the subsidiary questions without changing the answer to the primary question and with forward coding these subsidiary answers are lost.

2.9 For the 21-year data, the filter was checked both *forward* and *backward* in order to ensure that the loss of data was kept to a minimum. This is consistent with the approach adopted for the cleaning of data from the Ten-year Follow-up.

Table 2.1: Outline of 'correction' rules adopted for editing the 21-year data

Type of 'error'	Priority	'Correction' rules
Not valid	1	Check and correct any obvious data format problem (eg: left/right-shifted data)
	2	Set to missing
Out of range	1	Check and correct any obvious data format problem (eg: left/right-shifted data)
	2	Check if range appropriate, and modify and re-run edit if necessary
	3	Set to missing
Not consistent	1	Check and correct any obvious data format problem (eg: left/right-shifted data)
	2	Check filter structure, and modify and re-run edit if necessary
	3	Check and update value of primary filter if necessary and re-run edit
	4	Set to missing

Missing values

2.10 As part of the work to check and update the 21-year data, efforts have been made to standardise the allocation of missing values.

2.11 The standard missing value codes are as follows:

-3 Not stated
-8 Out of range

2.12 It is important to note that these codes are not always consistent with those allocated to other BCS70 datasets relating to the birth, 5, 10, 16, 26 and 30-year surveys. It is hoped that missing value codes will be rationalised when all BCS70 data are set up as a single database.

Derived variables

2.13 Although a number of derived variables have been created for the purposes of different projects using the 21-year data, we were only able to acquire the documentation of a limited number of derived variables, which have been included with the dataset deposited with the Data Archive.

2.14 Further details of the derived variables are given in Appendix 3.

INTERACTIVE DATA DICTIONARY

2.15 In order to help users find their way about the ***BCS70 Twenty one-year Sample Survey***, CLS has prepared an ***Interactive Data Dictionary*** as guide to the contents of the SPSS data. This is based on the ***Idealist*** Information Retrieval System, and is distributed, free of charge with the deposited data. It is available from the UK Data Archive and also from the Centre for Longitudinal Studies (appropriate url here or cohort@cls.ioe.ac.uk).

Contents

2.16 The ***Interactive Data Dictionary*** is supplied as two files:

- BCS70_21 Data Dictionary - the *Data Dictionary* in "ZIPped" form. This holds all the files which together make-up the Data Dictionary.

- README- notes on how to load the Data Dictionary on your PC.

YOU MAY COPY THESE FILES AS MANY TIMES AS YOU WISH.

IMPORTANT

2.17 Further details of the ***Interactive Data Dictionary*** are given below. Please read these carefully before attempting to load/use this software. If you have any problems or queries, please contact the User Support Group (see below for full contact information).

Loading the *Interactive Data Dictionary* from disk

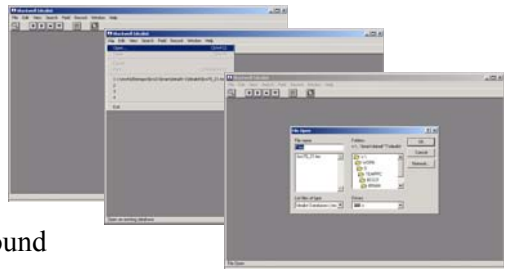
2.18 To load this documentation on your PC, follow the instructions given below. (See also file "READ.ME").

1. Users should extract the contents of the ' BCS70_21 Data Dictionary' ZIPfile into a separate directory

2. To use:

   - Within the Data Dictionary directory, click on 'Iwinread' – the Data Dictionary should load

   - If Idealist opens, but no database is loaded:
     - Within Idealist, click on 'File'
     - Within 'File', click on 'Open'
     - Within 'Open', select 'BCS_21.tex'



3. Opening screens will provide guidance on use and background on the survey (see also below).

4. Address queries to: cohort@cls.ioe.ac.uk

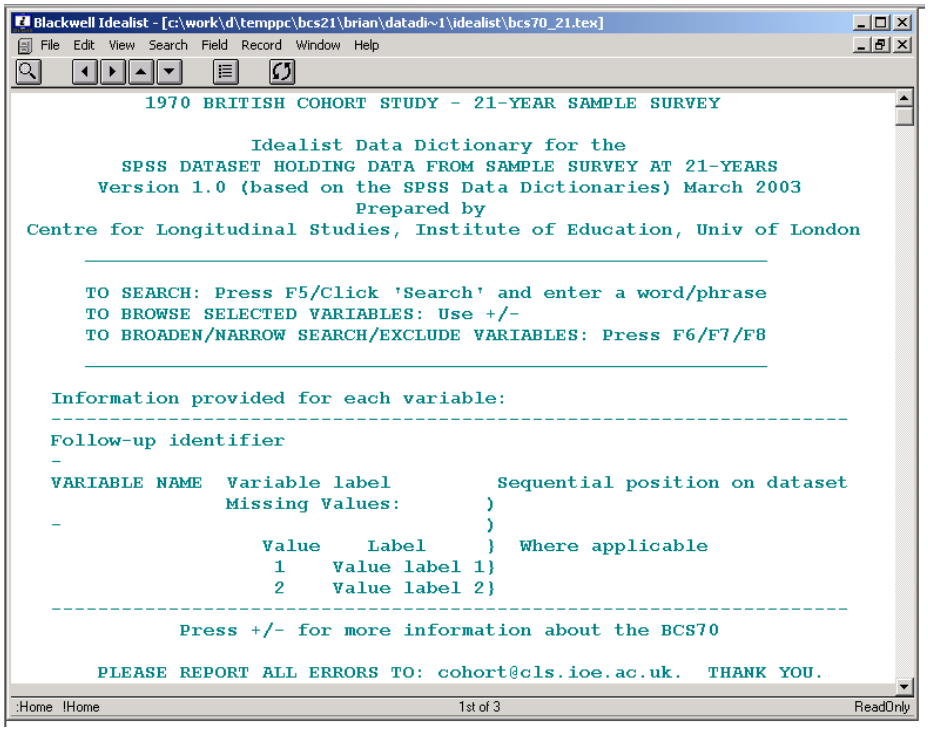
Guide to using the Interactive Data Dictionary

2.19 A brief guide to using the ***Interactive Data Dictionary*** is given below**.**

Starting the Data Dictionary

Within the Data Dictionary directory, click on 'Iwinread' to load the Data Dictionary.

You will be presented with an opening screen which includes a summary of the available search options and a guide to layout of the information, as well as access to more information about BCS70 and the 21-year Sample Survey:

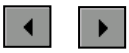


Search options

The Data Dictionary is designed to provide an interactive guide to the information held on the BCS70 21-year SPSS dataset, and is based on the SPSS 'data dictionary'.

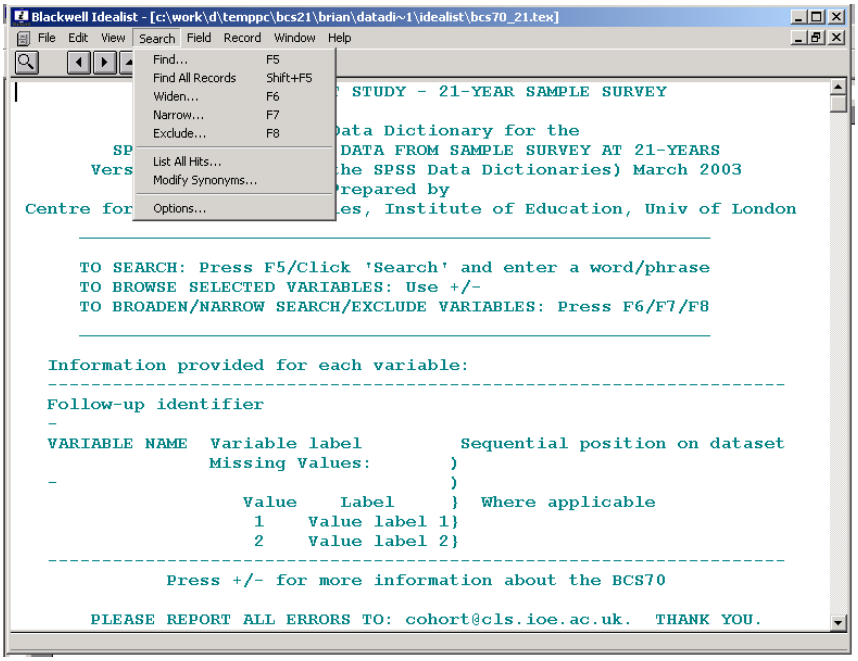
It provides the opportunity to search the variable names, variable labels and value labels for partuicular 'words' or 'phrases' and to generate a 'hit list' of all potentially relevant variables. To refine the search, the hit list can be broadened or narrowed, and selected variables may be excluded.

Searches my be initiated using the keyboard function keys; the drop down menus or menu buttons:

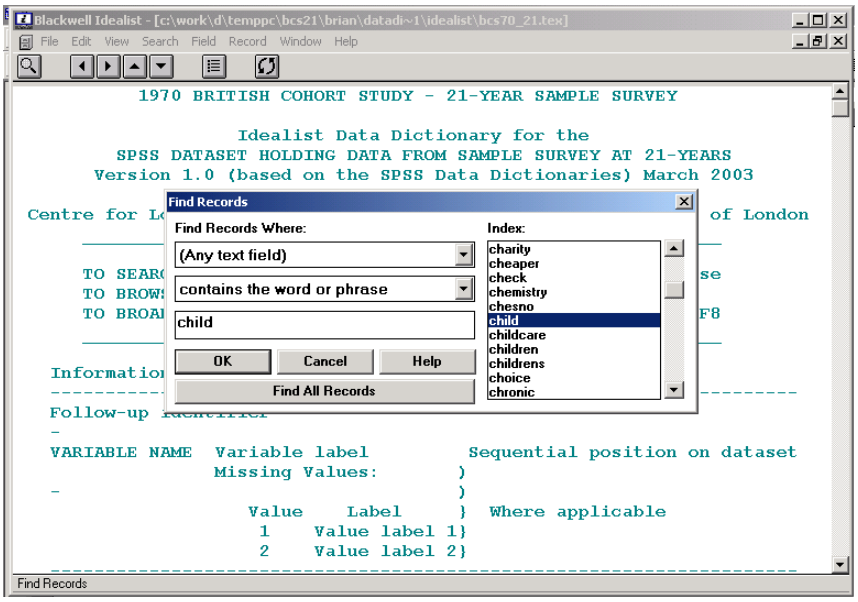
	Function key	**Menu**	**Button**
To SEARCH Use Function key/menu/button as follows:	F5	Search	
Enter	String/phrase to search for all examples of string/phrase in BCS70 21-year dataset variable names/variable labels/value labels		
NB:	A list of all indexed terms appears to the right of the dialog box. This may be browsed using cursor or by entering numbers/letters until derisred string is indicated.		
Click OK	This will select all records containing the chosen string.		
Hit List:	Selected records are known as a 'HIT LIST'		
To BROWSE hit list	+ and/or – keys on mumeric keypad	Not available	
To browse ENTRIES in hit list	Use the arrow/PgUp/PgDn/Home/End keys		
To WIDEN (add to) hit list	F6	Widen	Not available
To NARROW (shorten) hit list	F7	Narrow	Not available
To EXCLUDE (drop) from hit list	F8	Exclude	Not available

See examples below:

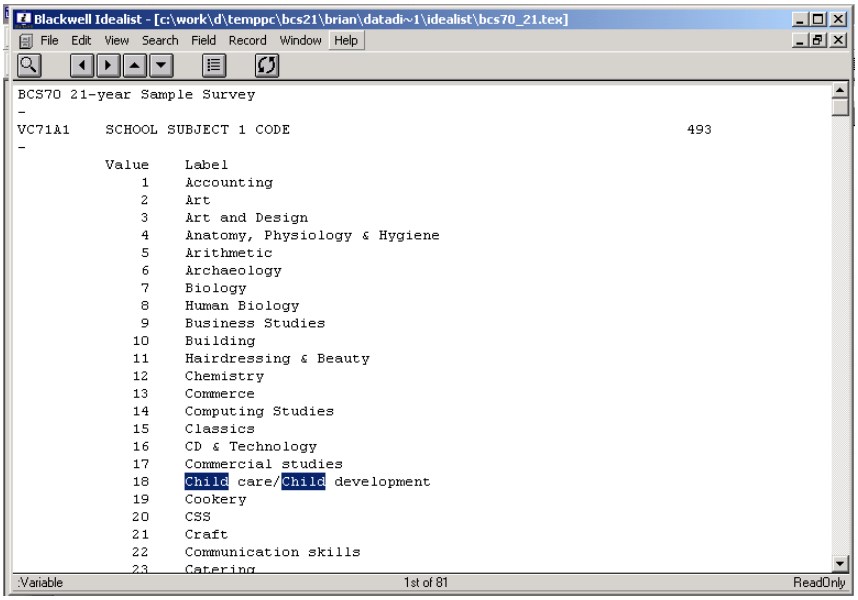
SEARCH Menu



SEARCH Dialog box for 'child'



Record #1 of the hit list for 'child'

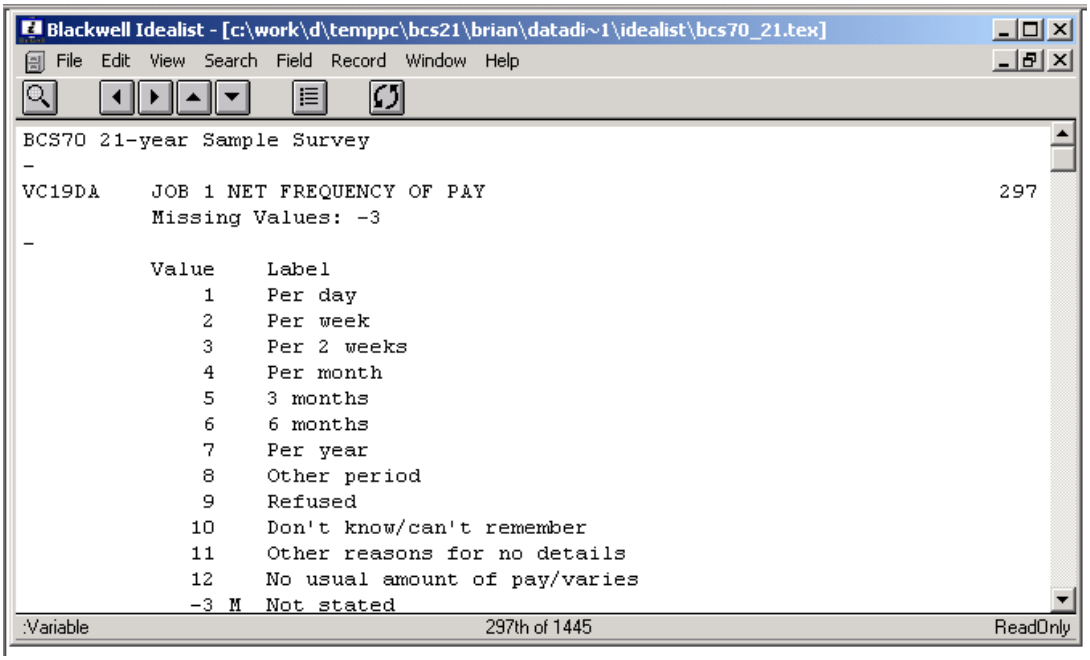


Layout of information

As the opening screen indicates, the information provided for each variable included in the 21-year dataset includes:

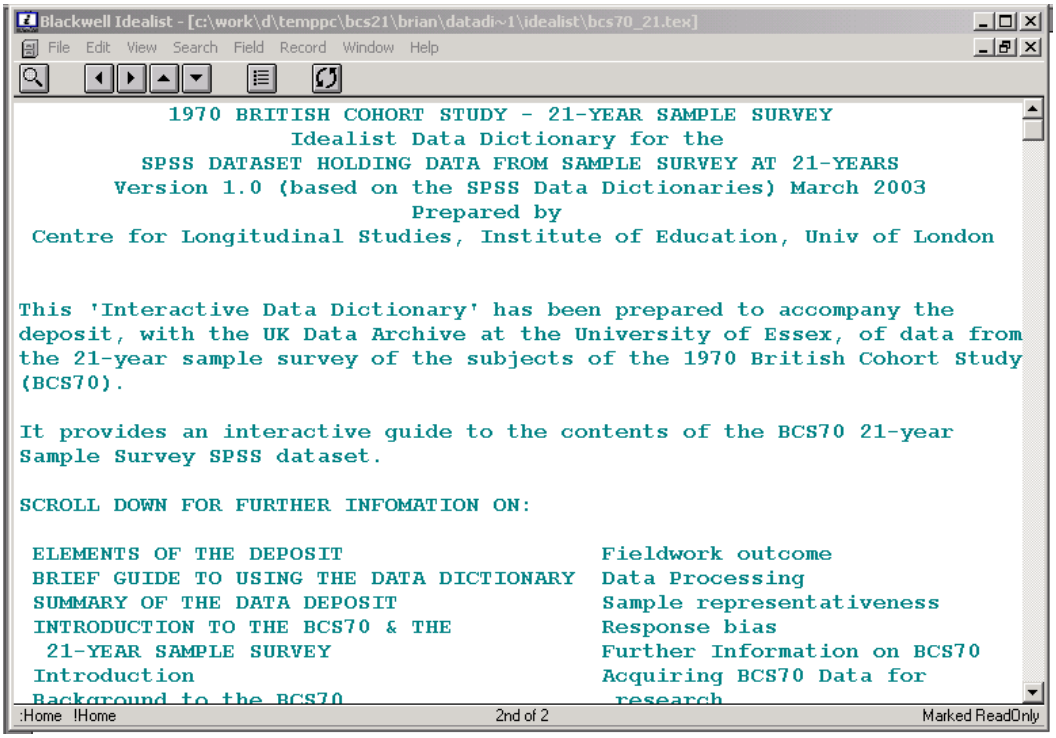
- Follow-up identifier (always 'BCS70 21-year Sample Survey')
- Variable name
- Variable label
- Missing values (where appropriate)
- Value labels (where appropriate)
- Sequential position in dataset

An example is given below:



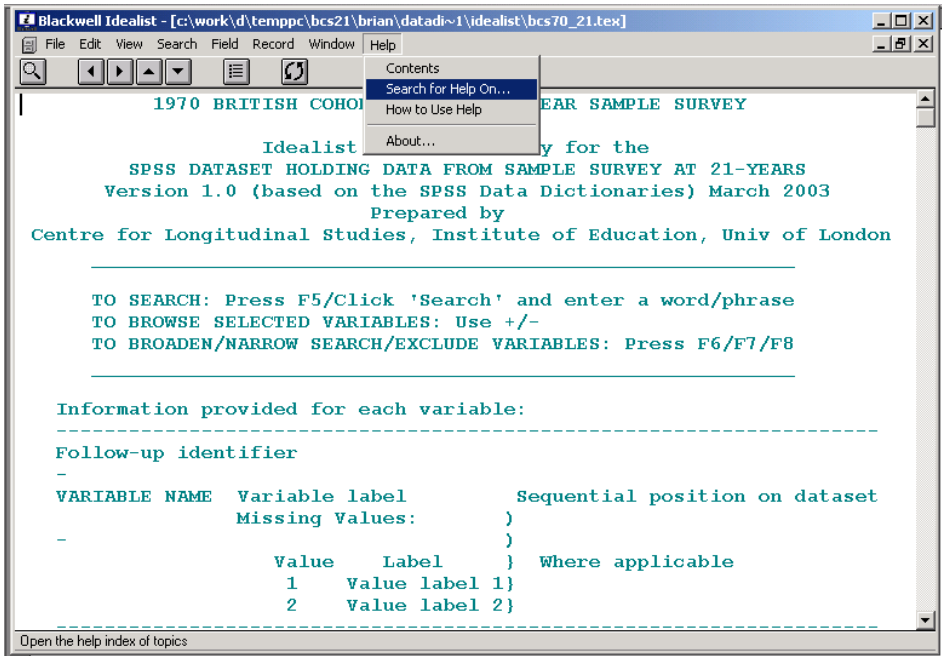
Information about BCS70 and the 21-year Sample Survey

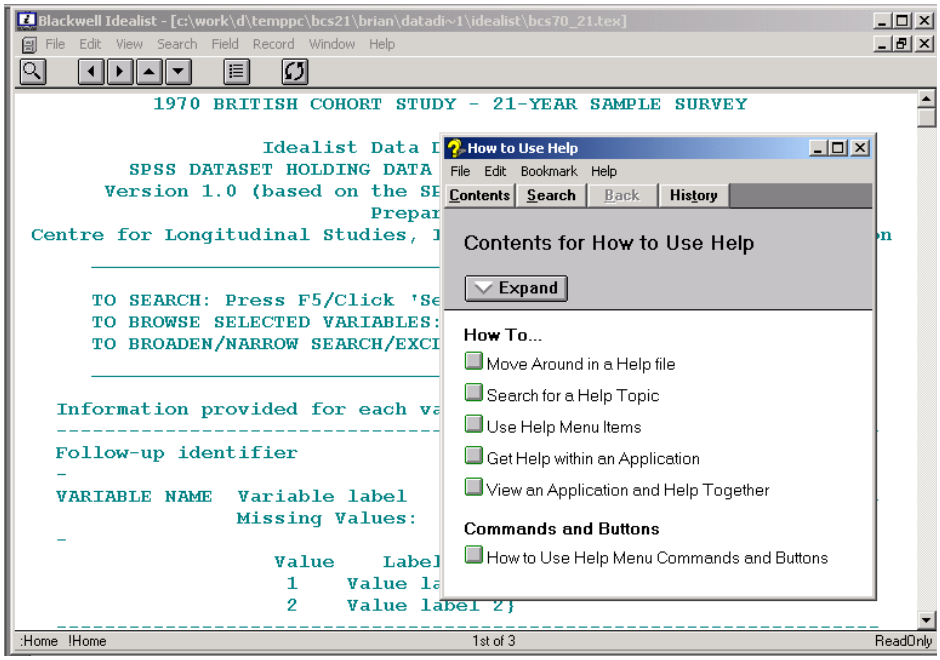
A brief overview of BCS70 in general, and the 21-year sample survey in partcular is presented after the opening screen and may be accessed by usning the + and/or – keys on mumeric keypad or the buttons.



Further information about ideaList

Further imformation about Idealist may be obtained using the built-in Help function which includes a useful guide on 'How to use help' – see below.





LONGITUDINAL LINKAGE TO EARLIER AND LATER DATASETS

2.20 The unique case identifier included with the BCS70 21-year data is the 6-digit variable KEY. The appropriate SPSS code for sorting cases and linking datasets would appear as follows:

```
sort cases by KEY.

match files file=filename1/
        file=filename2/
        by=KEY.
```

2.21 The population for the 1970 British Cohort Study is everyone living in Great Britain and born between 5-11th April 1970[1]. This population has grown between sweeps through sustained efforts to trace those missed at the birth survey, and also through immigration. These new cohort members were recruited for the follow-ups at ages 5, 10 and 16 years, and there were even a few additional cohort members apearing as late as the 26-year and 30-year follow-ups. As a result there are new KEY numbers appearing with each sweep. These cases will of course have no linkage to earlier datasets. For the *BCS70 Five-year Follow-up*, the new members to the study were given KEY numbers values in the 300,000s and 400,000s, and these cases will have no linkage to birth data. New members at the 10-year survey were given KEY values in the 600,000s or 700,000s, and those new at the 16-year survey or later, values in the 800,000s and 900,000s.

2.22 In December 2002, the CLS produced a document 'Cohort Studies Data Note 1: Longitudinal Linkage in BCS70', which can be downloaded from the CLS website. Users encountering problems in linking BCS70 datasets should contact the BCS70 User Support Group.

[1]The birth sweep covered the United Kingdom, but subsequent sweeps excluded Northern Ireland.