

THE ALSPAC STUDY

SATS FILE

SCHOOL TEST & EXAMINATION RESULTS DATA COLLECTED FROM LOCAL EDUCATION AUTHORITIES

Prepared by

**David Herrick
&
Jean Golding**

Documentation giving frequencies, background and instructions for use.

Last updated for version 1b of the built file.

December 2008

© University of Bristol

Introduction and Methodology

PLEASE NOTE:

The sample size you will receive from the education data is dependent on external factors including participant consent status and the data sharing agreement with the relevant organisation. These are subject to change without notice and cannot be controlled by ALSPAC. Every individual request for data is filtered against the current status at the time of data extraction, therefore if you request data at different time points it may be a different sample size each time.

February 2012

Key Stages & Tests

Under UK law, all children must be in school by their fifth birthday. In practice children will enter reception class in the September following their fourth birthday and then move up a school year every September. Their year on year progress is divided in to a number of “key stages”. See table 1 below for details (information obtained from the DfES website). Foundation, key stage 1 and key stage 2 make up primary education, with foundation / key stage 1 often referred to as “infants” and key stage 2 as “juniors”. Key stages 3 and 4 make up the compulsory component of secondary education.

At the end of each key stage, children’s progress is assessed by means of compulsory national tests. These are detailed in table 1 below. At the time when the ALSPAC cohort entered school there were no national compulsory entry assessments. However, the four LEAs that cover the former Avon area (Bristol, South Gloucestershire, Bath & North East Somerset and North Somerset) all employed the same entry assessment scheme, which was said to be used by 80% of the local state schools.

Table 1: Structure of UK schools and testing procedures

Class	Age	Key Stage	Tests
-	3 - 4	Foundation	-
Reception	4 - 5		Local entry assessments
Year 1	5 - 6	1	-
Year 2	6 - 7		National tests and tasks in English and Maths
Year 3	7 - 8	2	-
Year 4	8 - 9		-
Year 5	9 - 10		-
Year 6	10 - 11		National tests in English, Maths and Science
Year 7	11 - 12	3	-
Year 8	12 - 13		-
Year 9	13 - 14		National tests in English, Maths and Science
Year 10	14 - 15	4	Some children take GCSEs
Year 11	15 - 16		Most children take GCSEs, GNVQs etc.

Educational progress is measured by “national curriculum levels”. Eight levels are available, labelled Level 1 to Level 8. The way they fit into the key stages is detailed in table 2 below (information obtained from the DfES website). Note that the local entry assessments do not fit into this structure at all and that national curriculum levels are not used for assessment at key stage 4.

Table 2: National curriculum levels

Key Stage	Range of levels within which most children will work	Target that most children reach by the end of the key stage
1	1 - 3	2
2	2 - 5	4
3	3 - 7	5 - 6

Schooling of the ALSPAC cohort

The ALSPAC cohort is split across 3 school years. The expected progress, according to date of birth, of each subdivision of the cohort through compulsory schooling is shown in table 3. Note that the oldest ALSPAC children entered reception in autumn 1995 and the youngest will take their GCSEs in summer 2009. Table 3 also details the codes that will be used for the relevant school years on ALSPAC data files.

Table 3: Coding of school years on ALSPAC data files and expected progress of the ALSPAC cohort according to their dates of birth

School year	Code	Expected position of ALSPAC children born between:		
		April 1991 & August 1991	September 1991 & August 1992	September 1992 & January 1993
1995 / 1996	1	Reception	-	-
1996 / 1997	2	Year 1	Reception	-
1997 / 1998	3	Year 2	Year 1	Reception
1998 / 1999	4	Year 3	Year 2	Year 1
1999 / 2000	5	Year 4	Year 3	Year 2
2000 / 2001	6	Year 5	Year 4	Year 3
2001 / 2002	7	Year 6	Year 5	Year 4
2002 / 2003	8	Year 7	Year 6	Year 5
2003 / 2004	9	Year 8	Year 7	Year 6
2004 / 2005	10	Year 9	Year 8	Year 7
2005 / 2006	11	Year 10	Year 9	Year 8
2006 / 2007	12	Year 11	Year 10	Year 9
2007 / 2008	13	-	Year 11	Year 10
2008 / 2009	14	-	-	Year 11

Data Source

This built file currently contains data from school entry assessments and key stage 1 assessments. The four LEAs that cover the former Avon area (Bristol, South Gloucestershire, Bath & North East Somerset and North Somerset) were approached and agreed that ALSPAC would be allowed to access these data. The data were actually provided by South Gloucestershire, as they hold data for the other LEAs. The data received consisted of entry assessment results and key stage 1 results from the school years indicated in table 3 for children born between 01/04/91 – 31/12/92.

However, Bristol LEA also sent a data set. The data they provided were actually complete data for the relevant school years, with no birth date restriction. This allowed records for any premature or late ALSPAC deliveries who did their entry assessments or key stage 1 results in schools within the Bristol LEA to be identified. Their data were added to the main dataset provided by South Gloucestershire.

Unfortunately the source files contain only a school code, which is only unique within LEAs, and no code for the LEA. Comparison of the Bristol only data with the data provided by South Gloucestershire indicates that there are some school codes that are repeated within different LEAs. It is not possible to unpick this to uniquely identify which school each child is in, so school identifiers are not provided in this data set.

Permissions

Permission to link to general ALSPAC data was originally sought from the accompanying adult at the Focus@7 clinic. However, in May 2003 the ALSPAC Ethics & Law Committee decided that in line with the data protection act it was not necessary for ALSPAC to obtain written consent before using the data, but that data on children for whom permission was actually refused at the Focus@7 clinic or for whom permission was subsequently withdrawn should not be used. At the time of writing there are 15 children for whom we have such refusals.

Linking

The entry assessment and key stage 1 records were first linked to each other using the LEA PIN numbers and name, date of birth and sex. Some corrections to the PIN numbers were required. The resulting master list of children was then linked to ALSPAC identifiers using names and dates of birth. This linking was done in tandem with linking the SA file data by first linking to a list of all known deliveries with an identifier called "joinedchild".

Identifier Variables

Records on this file are identified by ALN & QLET (the standard ALSPAC child identifiers).

Multiple Records

Multiple records were returned for a small number of children. These occurred for one of four reasons:

1. The child moved schools during Year 3. This means there are two records in the same sweep but from different schools.
2. The child repeated Year 3. This means there are two records from the same school but in different sweeps.
3. The child did Year 3 twice in different schools. This means there are two records from different schools in different sweeps.

One record was selected for each child as follows. For 2 & 3 above the record that came from the year in which the child should have sat the tests by age (see tables 3 and variables SAT006a & SAT106a) was retained. For 1 above the most complete record was retained. If it was not possible to distinguish on the basis of fewest missing values then the less good set of results was retained.

Sample

13,380 entry assessment records and 15,861 key stage 1 records were identified at the linking stage as belonging to children eligible for ALSPAC. 59 pairs of multiple records were present in the entry assessment data and 52 in the key stage 1 data. Selection of a single record for each of these children left, therefore, 13,321 entry assessment records and 15,809 key stage 1 records for this built file.

The sample of children with entry assessment records is not a subset of those with key stage 1 records. A total of 16,367 children are present on the combined sample on which this built file is based. Note that some of these will relate to the “new cases” first invited to participate in ALSPAC at Focus@7. Please refer to §2.4 of the Focus@7 documentation for more information about new cases.

Of the 13,321 entry assessment records, 9,482 belong to children in the core ALSPAC sample (14,676 fetuses, 14,062 live births). The remaining 3,839 records belong to eligible children not in the core sample. For reasons of confidentiality 6 records belonging to children from triplet and quadruplet pregnancies are not available. All of these 6 children are in the core ALSAPC sample.

Of the 15,809 key stage 1 records, 11,279 belong to children in the core ALSPAC sample. The remaining 4,530 records belong to eligible children not in the core sample. For reasons of confidentiality 13 records belonging to children from triplet and quadruplet pregnancies are not available. 10 of these 13 children are in the core ALSPAC sample.

Version History

Version 1a – April 2004

The first version of the SATS built file, containing entry assessment and key stage 1 data.

Version 1b – December 2008

5 records from duplicate enrolments that were identified as carrying the ALN that was scrapped had ALN recoded to the correct retained value for that pregnancy. All 5 had records for both entry assessment and key stage 1. 25 records identified as belonging to ineligible pregnancies were removed. All 25 had records for key stage 1, but only 15 had records for entry assessment.

Due to the extension of direct access to ALSPAC data to non-ALSPAC staff and in order to comply with guidance issued in 1996 by the ALSPAC Law & Ethics Committee regarding the confidentiality of multiple pregnancies variable SAT001 has been recoded to 2 “No” for 6 children from triplet and quadruplet pregnancies for whom entry assessment records were identified and all other entry assessment variables set to -11. Similarly, SAT101 has been recoded to 2 “No” for 13 children from triplet and quadruplet pregnancies for whom key stage 1 records were identified and all other key stage 1 variables set to -11. Note that 3 of these 13 cases are not in the core sample.

References

1. South Gloucestershire Professional & Curriculum Support Service. 1996 *Entry Assessment in South Gloucestershire*. Weston-super-mare: Woodspring Educational Resource Centre. (ISBN 1901084027, one copy held by ALSPAC)
2. Qualifications and Curriculum Authority. 1998. *Key Stage 1 1999 - Assessment and reporting arrangements*. (Copy held in “School Assessment Folder - Key Stage 1 1999” in University of Bristol Education Library.)

Entry Assessment

General Comments

The following information was obtained from the booklet “Entry Assessment in South Gloucestershire” (1):

Entry Assessment observes eight areas of experience. Four areas form the required element (language, reading, writing and mathematics) and the other four are voluntary (social skills, problem solving, large motor skills and small motor skills). The stage reached by each child for each area (one of 2 to 7, with 7 being the most advanced and 2 the least advanced) should be recorded, or A for “child absent” and U for “child unassessed”.

In the data received by ALSPAC, codes A and U only appeared in the 1997 / 1998 data. Code X was also present in the data for that year, but only in the data for the voluntary areas from the LEAs other than Bristol. The data for 1995 / 1996 and 1996 / 1997 contained a code 0, which presumably corresponds to the official codes A and U. These inconsistencies in the source data mean that it is not possible to disentangle the different reasons for an absence of a score. Codes A, U, X and 0 were therefore simply all set to -1.

Administrative Variables

Data available for child

An indicator of whether entry assessment data was supplied for each child is available as SAT001:

sat001 Entry assessment: Data available for child

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	13315	81.4	81.4	81.4
	2 No	3052	18.6	18.6	100.0
	Total	16367	100.0	100.0	

If SAT001 = 2 then all entry assessment variables were set to -6.

School year taken

This is the school year in which the child sat their entry assessments:

sat006 Entry assessment: School year taken

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 1995 / 1996	3229	24.3	24.3	24.3
	2 1996 / 1997	7276	54.6	54.6	78.9
	3 1997 / 1998	2810	21.1	21.1	100.0
	Total	13315	100.0	100.0	

Child is in correct school year

An indicator of whether the child sat the entry assessments in the correct year according to their date of birth (see table 1) was derived as SAT006a:

sat006a Entry assessment: Child is in correct school year

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	13291	99.8	99.8	99.8
	2 No	24	.2	.2	100.0
	Total	13315	100.0	100.0	

Required Areas

sat010 Entry assessment: Language

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	44	.3	.3	.3
	3	237	1.8	1.8	2.1
	4	2469	18.5	18.6	20.7
	5	4239	31.8	31.9	52.6
	6	3236	24.3	24.3	76.9
	7	3071	23.1	23.1	100.0
	Total	13296	99.9	100.0	
Missing	-1 Missing	19	.1		
Total		13315	100.0		

sat020 Entry assessment: Reading

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	138	1.0	1.0	1.0
	3	357	2.7	2.7	3.7
	4	2896	21.7	21.8	25.5
	5	6561	49.3	49.3	74.9
	6	2803	21.1	21.1	95.9
	7	540	4.1	4.1	100.0
	Total	13295	99.8	100.0	
Missing	-1 Missing	20	.2		
Total		13315	100.0		

sat030 Entry assessment: Writing

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	145	1.1	1.1	1.1
	3	530	4.0	4.0	5.1
	4	3978	29.9	29.9	35.0
	5	5383	40.4	40.5	75.5
	6	3228	24.2	24.3	99.7
	7	35	.3	.3	100.0
	Total	13299	99.9	100.0	
Missing	-1 Missing	16	.1		
Total		13315	100.0		

sat040 Entry assessment: Mathematics

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	128	1.0	1.0	1.0
	3	454	3.4	3.4	4.4
	4	3043	22.9	22.9	27.3
	5	4959	37.2	37.3	64.6
	6	2844	21.4	21.4	86.0
	7	1867	14.0	14.0	100.0
	Total	13295	99.8	100.0	
Missing	-1 Missing	20	.2		
Total		13315	100.0		

Voluntary Areas

sat050 Entry assessment: Social skills

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	103	.8	1.4	1.4
	3	250	1.9	3.5	4.9
	4	887	6.7	12.4	17.3
	5	2815	21.1	39.4	56.7
	6	1966	14.8	27.5	84.2
	7	1132	8.5	15.8	100.0
	Total	7153	53.7	100.0	
Missing	-1 Missing	6162	46.3		
Total		13315	100.0		

It should also be noted that the uptake of the social skills test was very high in 1997 / 1998 so it would appear from the data this test was compulsory in this year. There were two cases with codes that did not make sense for this variable. One was coded 1 and the other 9. These two cases were recoded to -1 "Missing".

sat060 Entry assessment: Problem solving

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	9	.1	.2	.2
	3	53	.4	1.3	1.5
	4	769	5.8	19.0	20.5
	5	1477	11.1	36.5	57.0
	6	1119	8.4	27.6	84.6
	7	623	4.7	15.4	100.0
	Total	4050	30.4	100.0	
Missing	-1 Missing	9265	69.6		
Total		13315	100.0		

sat070 Entry assessment: Large motor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	12	.1	.3	.3
	3	49	.4	1.1	1.3
	4	1166	8.8	25.6	26.9
	5	1247	9.4	27.3	54.2
	6	1553	11.7	34.0	88.3
	7	536	4.0	11.7	100.0
	Total	4563	34.3	100.0	
Missing	-1 Missing	8752	65.7		
Total		13315	100.0		

sat080 Entry assessment: Small motor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	14	.1	.3	.3
	3	63	.5	1.2	1.5
	4	1461	11.0	28.7	30.2
	5	1132	8.5	22.2	52.5
	6	1302	9.8	25.6	78.1
	7	1116	8.4	21.9	100.0
	Total	5088	38.2	100.0	
Missing	-1 Missing	8227	61.8		
Total		13315	100.0		

Derived Variables

Three summary scores have been derived. The most useful of these will be the required areas score (SAT090a/b/c) as this is available for nearly all the children for whom we have entry assessment data. The voluntary areas and all areas scores are dependent on the children having completed the voluntary areas.

Required areas score

This is the sum of the scores obtained in the four required areas, translated so that the minimum possible score is 0. Two versions of the score are provided. For most analyses the prorated score, SAT090b, should be used.

The required areas score was created by recoding (missing, 2 = 0) (3 = 1) (4 = 2) (5 = 3) (6 = 4) (7 = 5) in SAT010 to SAT040 and summing. The score based on complete cases only is SAT090a. If any of the components were missing then SAT090a was set to -5. Variable SAT090c is the number of missing components. The prorated score is SAT090b. If more than two components were missing then SAT090b was set to -5. Otherwise SAT090b was scaled by a factor of $4 / (4 - \text{SAT090c})$ and rounded to the nearest integer.

sat090a Entry assessment: Required areas score (complete cases)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	10	.1	.1	.1
	1	15	.1	.1	.2
	2	26	.2	.2	.4
	3	39	.3	.3	.7
	4	67	.5	.5	1.2
	5	102	.8	.8	1.9
	6	183	1.4	1.4	3.3
	7	333	2.5	2.5	5.8
	8	720	5.4	5.4	11.3
	9	949	7.1	7.1	18.4
	10	1211	9.1	9.1	27.5
	11	1468	11.0	11.1	38.6
	12	1550	11.6	11.7	50.2
	13	1525	11.5	11.5	61.7
	14	1354	10.2	10.2	71.9
	15	1160	8.7	8.7	80.6
	16	950	7.1	7.2	87.8
	17	767	5.8	5.8	93.6
	18	587	4.4	4.4	98.0
	19	253	1.9	1.9	99.9
	20	16	.1	.1	100.0
	Total	13285	99.8	100.0	
Missing	-5 1+ components missing	30	.2		
Total		13315	100.0		

sat090b Entry assessment: Required areas score (prorated)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	10	.1	.1	.1
	1	15	.1	.1	.2
	2	26	.2	.2	.4
	3	39	.3	.3	.7
	4	67	.5	.5	1.2
	5	102	.8	.8	1.9
	6	185	1.4	1.4	3.3
	7	334	2.5	2.5	5.8
	8	722	5.4	5.4	11.3
	9	950	7.1	7.1	18.4
	10	1211	9.1	9.1	27.5
	11	1471	11.0	11.1	38.6
	12	1552	11.7	11.7	50.3
	13	1527	11.5	11.5	61.7
	14	1356	10.2	10.2	71.9
	15	1160	8.7	8.7	80.6
	16	951	7.1	7.1	87.8
	17	767	5.8	5.8	93.6
	18	587	4.4	4.4	98.0
	19	253	1.9	1.9	99.9
	20	16	.1	.1	100.0
	Total	13301	99.9	100.0	
Missing	-5 > 2 components missing	14	.1		
Total		13315	100.0		

sat090c Entry assessment: Number of missing required areas components

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	13285	99.8	99.8	99.8
	1	11	.1	.1	99.9
	2	5	.0	.0	99.9
	3	2	.0	.0	99.9
	4	12	.1	.1	100.0
	Total	13315	100.0	100.0	

Voluntary areas score

This is the sum of the scores obtained in the four voluntary areas, translated so that the minimum possible score is 0. Two versions of the score are provided. For most analyses the prorated score, SAT091b, should be used.

The voluntary areas score was created by recoding (missing, 2 = 0) (3 = 1) (4 = 2) (5 = 3) (6 = 4) (7 = 5) in SAT050 to SAT080 and summing. The score based on complete cases only is SAT091a. If any of the components were missing then SAT091a was set to -5. Variable SAT091c is the number of missing components. The prorated score is SAT091b. If more than two components were missing then SAT091b was set to -5. Otherwise SAT091b was scaled by a factor of $4 / (4 - \text{SAT091c})$ and rounded to the nearest integer.

sat091a Entry assessment: Voluntary areas score (complete cases)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	2	.0	.1	.1
	1	1	.0	.0	.1
	2	2	.0	.1	.1
	3	3	.0	.1	.2
	4	8	.1	.2	.4
	5	14	.1	.4	.8
	6	29	.2	.8	1.6
	7	48	.4	1.3	2.9
	8	138	1.0	3.7	6.6
	9	214	1.6	5.8	12.4
	10	314	2.4	8.5	20.9
	11	359	2.7	9.7	30.7
	12	380	2.9	10.3	41.0
	13	377	2.8	10.2	51.2
	14	389	2.9	10.5	61.7
	15	404	3.0	10.9	72.7
	16	340	2.6	9.2	81.9
	17	263	2.0	7.1	89.0
	18	220	1.7	6.0	95.0
	19	128	1.0	3.5	98.4
	20	58	.4	1.6	100.0
Total		3691	27.7	100.0	
Missing	-5 1+ components missing	9624	72.3		
Total		13315	100.0		

sat091b Entry assessment: Voluntary areas score (prorated)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5	.0	.1	.1
	1	1	.0	.0	.1
	2	3	.0	.1	.2
	3	3	.0	.1	.2
	4	14	.1	.3	.5
	5	21	.2	.4	.9
	6	41	.3	.8	1.7
	7	61	.5	1.2	2.9
	8	265	2.0	5.2	8.2
	9	262	2.0	5.2	13.4
	10	431	3.2	8.5	21.9
	11	423	3.2	8.4	30.3
	12	562	4.2	11.1	41.4
	13	451	3.4	8.9	50.3
	14	513	3.9	10.1	60.4
	15	477	3.6	9.4	69.9
	16	568	4.3	11.2	81.1
	17	333	2.5	6.6	87.7
	18	331	2.5	6.5	94.2
	19	168	1.3	3.3	97.5
	20	124	.9	2.5	100.0
	Total	5057	38.0	100.0	
Missing	-5 > 2 components missing	8258	62.0		
Total		13315	100.0		

sat091c Entry assessment: Number of missing voluntary areas components

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3691	27.7	27.7	27.7
	1	595	4.5	4.5	32.2
	2	771	5.8	5.8	38.0
	3	2763	20.8	20.8	58.7
	4	5495	41.3	41.3	100.0
	Total	13315	100.0	100.0	

All areas score

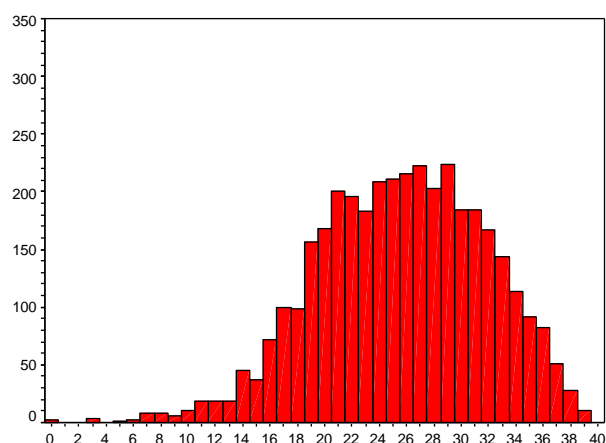
This is the sum of the scores obtained in all eight areas, translated so that the minimum possible score is 0. Two versions of the score are provided. For most analyses the prorated score, SAT092b, should be used.

The all areas score was created by recoding (missing, 2 = 0) (3 = 1) (4 = 2) (5 = 3) (6 = 4) (7 = 5) in SAT010 to SAT080 and summing. The score based on complete cases only is SAT092a. If any of the components were missing then SAT092a was set to -5. Variable SAT092c is the number of missing components. The prorated score is SAT092b. This was calculated by summing SAT090b and SAT091b, so that it is set to -5 if more than 2 of the required areas are missing or if more than two of the voluntary areas are missing.

Statistics

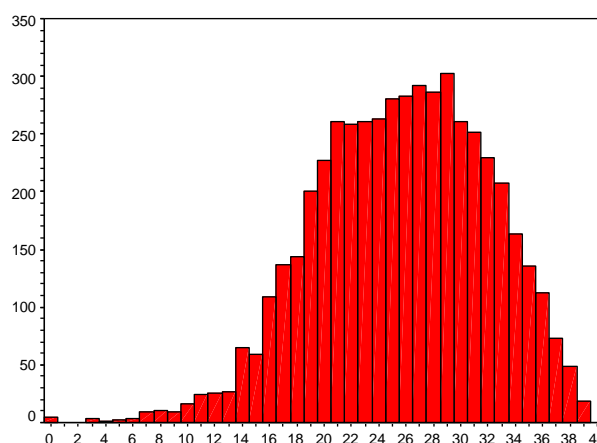
		sat092a Entry assessment: All areas score (complete cases)	sat092b Entry assessment: All areas score (prorated)
N	Valid	3686	5054
	Missing	12681	11313
Mean		25.64	25.68
Mode		29	29
Std. Deviation		6.149	6.282
Minimum		0	0
Maximum		39	39
Percentiles	25	21.00	21.00
	50	26.00	26.00
	75	30.00	30.00

SAT092a



Entry assessment: All areas score (complete cases)

SAT092b



Entry assessment: All areas score (prorated)

sat092c Entry assessment: Number of missing all areas components

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3686	27.7	27.7	27.7
	1	596	4.5	4.5	32.2
	2	772	5.8	5.8	38.0
	3	2760	20.7	20.7	58.7
	4	5480	41.2	41.2	99.8
	5	5	.0	.0	99.9
	6	6	.0	.0	99.9
	8	10	.1	.1	100.0
	Total	13315	100.0	100.0	

Age at testing

The child's age at testing was calculated in both weeks and months. Dates of birth were taken from central ALSPAC records. There is no specific date of testing in the source data (it was most likely not recorded). According to the test booklet, the test should have been completed within the first half of the autumn term and should have been initiated as soon as the teacher was satisfied that the children were settled. Based on this, the date of testing is assumed to be October 1st in the relevant years.

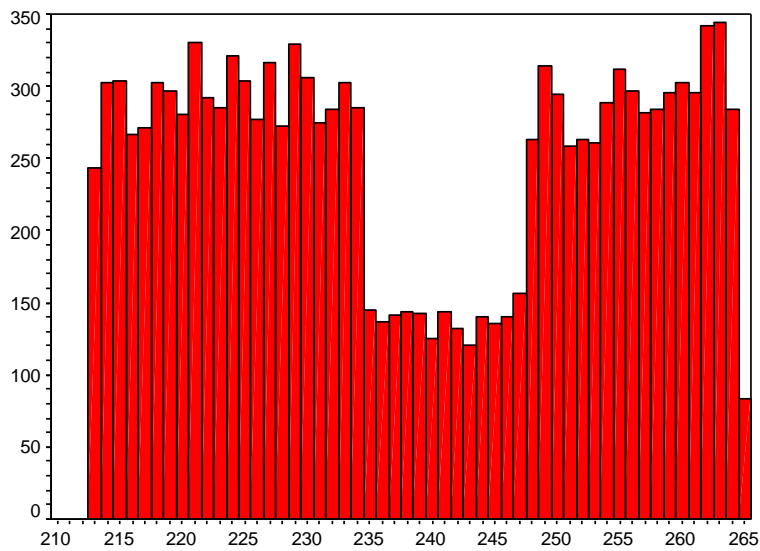
sat095a Entry assessment: Age of child at testing (months)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	44	1	.0	.0	.0
	45	1	.0	.0	.0
	49	1218	9.1	9.1	9.2
	50	1336	10.0	10.0	19.2
	51	1274	9.6	9.6	28.8
	52	1315	9.9	9.9	38.6
	53	1253	9.4	9.4	48.1
	54	651	4.9	4.9	52.9
	55	549	4.1	4.1	57.1
	56	590	4.4	4.4	61.5
	57	1254	9.4	9.4	70.9
	58	1211	9.1	9.1	80.0
	59	1293	9.7	9.7	89.7
	60	1332	10.0	10.0	99.7
	61	24	.2	.2	99.9
	62	2	.0	.0	99.9
	63	4	.0	.0	99.9
	64	1	.0	.0	100.0
	65	4	.0	.0	100.0
	67	1	.0	.0	100.0
	69	1	.0	.0	100.0
Total		13315	100.0	100.0	

Age in weeks (SAT095b) is presented separately as a histogram for the children in the correct school year (SAT006a = 1) and as a frequency table for those not in the correct school year (SAT006a = 2).

Children in correct school year:

SAT095B



Entry assessment: Age of child at testing (weeks)

Children not in correct school year (i.e. outliers):

sat095b Entry assessment: Age of child at testing (weeks)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	193	4.2	4.2	4.2
	198	4.2	4.2	8.3
	265	8.3	8.3	16.7
	266	8.3	8.3	25.0
	267	12.5	12.5	37.5
	268	8.3	8.3	45.8
	271	4.2	4.2	50.0
	272	4.2	4.2	54.2
	274	8.3	8.3	62.5
	276	4.2	4.2	66.7
	277	4.2	4.2	70.8
	279	4.2	4.2	75.0
	283	4.2	4.2	79.2
	284	12.5	12.5	91.7
	293	4.2	4.2	95.8
	301	4.2	4.2	100.0
Total	24	100.0	100.0	

Key Stage 1

General Comments

The scoring system for Key Stage 1 is somewhat more complicated than that for Entry Assessment. The following information comes from the DfEE booklet “Key Stage 1 1999 - Assessment and reporting arrangements” (2). Key Stage 1 consists of teacher assessments and standardised tasks and tests. The teacher assessments record a level of achievement for speaking and listening, reading, writing, mathematics and science. The standardised tasks and tests consist of a reading task, a reading comprehension test, a writing task, a spelling test and mathematics. Although results of the standardised tasks and tests were obtained for all three school years relevant to the ALSPAC cohort, data on the teacher assessments were only obtained for the third school year (1999 / 2000). The following information, therefore, concentrates on the standardised tasks and tests.

The basic scale consists of levels 1, 2, 3 and 4+, with grades A, B and C within level 2. Level 4+ is assessed by means of Key Stage 2 materials. There were so few children achieving level 4+ that it has been combined with level 3. This also makes the results a simple measure of achievement at Key Stage 1 and is fairer as the option of being tested beyond level 3 would not have been available to all capable children. In addition, code W (“working towards level 1”) means that the child was assessed but didn’t achieve level 1, code A means the child was absent and code D that the child was “disapplied”, e.g. due to special educational needs. However, not all levels are available for all tests.

The data for 1997 / 1998 and 1998 / 1999 were received with the gradings within level 2 stored as separate variables. These were combined with the relevant variables to produce a scale of W, 1, 2C, 2B, 2A, 3, 4+. The data for 1999 / 2000 were presented as text variables in this way anyway. On the built file the data are presented as integer variables, with W = 0, 1 = 1, 2C = 2, 2B = 3, 2A = 4, 3 / 4+ = 5.

Administrative Variables

Data available for child

An indicator of whether key stage 1 data was supplied for each child is available as SAT101:

sat101 Key stage 1: Data available for child

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Yes	15796	96.5	96.5	96.5
2 No	571	3.5	3.5	100.0
Total	16367	100.0	100.0	

If SAT101 = 2 then all key stage 1 variables were set to -6.

School year taken

This is the school year in which the child sat their key stage 1 assessments:

sat106 Key stage 1: School year taken

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 3 1997 / 1998	3848	24.4	24.4	24.4
4 1998 / 1999	8996	57.0	57.0	81.3
5 1999 / 2000	2952	18.7	18.7	100.0
Total	15796	100.0	100.0	

Child is in correct school year

An indicator of whether the child sat the key stage 1 assessments in the correct year according to their date of birth (see table 1) was derived as SAT106a:

sat106a Key stage 1: Child is in correct school year

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Yes	15757	99.8	99.8	99.8
2 No	39	.2	.2	100.0
Total	15796	100.0	100.0	

Reading Task and Reading Comprehension Test

These form a dual assessment of reading. The instructions for administering them are as follows:

- The reading task awards levels 1 and 2 only (with grades C - A at level 2).
- Those children achieving level 2 in the reading task are then entered for the reading comprehension test.
- The reading comprehension test first awards level 2 (with grades C - A).
- Children achieving grade A at level 2 in both the reading task and reading comprehension test are then assessed with a level 3 reading comprehension test.

The instructions for recording the results of the reading task and the reading comprehension test are as follows:

- Children scoring W or 1 for the reading task are marked as X ("not required to be entered for the test") for the reading comprehension test.
- Children who do not achieve level 2 on the reading comprehension test have recorded their reading task mark and L ("lower than level 2 threshold") for the reading comprehension test.
- For children achieving level 2 on both the reading task and reading comprehension test both scores are recorded.
- For children achieving level 3 on the reading comprehension test, only this score is recorded and the level of achievement on the reading task is omitted. (Note that it must have been 2A.)
- If a child achieves level 4 or above for reading via Key Stage 2 materials, level 4+ is entered for the reading comprehension test and, again, the reading task level is omitted. Note that one case had 5 recorded for the reading comprehension test. Presumably the child genuinely reached level 5 on the Key Stage 2 materials but the exact level was recorded in error instead of 4+. This is therefore treated the same as 4+ and combined with level 3.

The above rubric led to some logical edits being performed on the raw data as follow:

- 20 cases had 2A recorded for the reading task as well as 3 or 4+ for the reading comprehension test. These reading task results were recoded (along with all the remaining missing values, which all had scores of 3 or 4+ for the reading comprehension test) to -1 "RCT = 3/4+".
- One case had 2B recorded for the reading task and 3 for the reading comprehension test. From the booklet it is clear that the reading comprehension tests might have been carried out before the reading task was completed. In these circumstances the child should have taken the reading comprehension test if expected to achieve level 2 in the reading task. It seems likely that this is what happened in this case and that the child achieved 2A in the reading comprehension test and so went on to do the level 3 test, but subsequently only achieved 2B in the reading task. The 3 for the reading comprehension test for these cases has therefore been recoded to 2A.
- The final edits concern cases marked A or D for the reading task. Whilst most of these have the same code for the reading comprehension test, some are marked as X and 3 of those marked A for the reading task have reached level 2 in the reading comprehension test. Since the absence of a mark for the reading task should mean the pupils are not entered for the reading comprehension test, all pupils with A or D for

the reading task were coded X for the reading comprehension test. Note, however, that there are 19 cases absent for the reading comprehension test that achieved level 2 for the reading task. These cases have been left in, as they do have some meaning.

The resulting variables are:

sat120 Key stage 1: Reading task

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid				
0 W	496	3.1	4.4	4.4
1 1	2444	15.5	21.7	26.1
2 2C	2668	16.9	23.7	49.8
3 2B	3186	20.2	28.3	78.1
4 2A	2472	15.6	21.9	100.0
Total	11266	71.3	100.0	
Missing				
-5 Absent (A)	36	.2		
-4 Disapplied (D)	61	.4		
-1 RCT = 3/4+	4433	28.1		
Total	4530	28.7		
Total	15796	100.0		

sat125 Key stage 1: Reading comprehension test

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid				
2 2C	2450	15.5	19.9	19.9
3 2B	2676	16.9	21.7	41.6
4 2A	2761	17.5	22.4	64.0
5 3 / 4+	4433	28.1	36.0	100.0
Total	12320	78.0	100.0	
Missing				
-5 Absent (A)	21	.1		
-3 < level 2 (L)	418	2.6		
-2 Not entered (X)	3037	19.2		
Total	3476	22.0		
Total	15796	100.0		

As an aid to understanding the rubric above, the cross-tabulation of these two variables may be useful:

sat120 Key stage 1: Reading task * sat125 Key stage 1: Reading comprehension test Crosstabulation

Count		sat125 Key stage 1: Reading comprehension test							Total
		-5 Absent (A)	-3 < level 2 (L)	-2 Not entered (X)	2 2C	3 2B	4 2A	5 3 / 4+	
sat120	-5 Absent (A)	0	0	36	0	0	0	0	36
Key stage	-4 Disapplied (D)	0	0	61	0	0	0	0	61
1: Reading	-1 RCT = 3/4+	0	0	0	0	0	0	4433	4433
task	0 W	0	0	496	0	0	0	0	496
	1 1	0	0	2444	0	0	0	0	2444
	2 2C	14	389	0	1659	485	121	0	2668
	3 2B	5	26	0	678	1469	1008	0	3186
	4 2A	2	3	0	113	722	1632	0	2472
Total		21	418	3037	2450	2676	2761	4433	15796

Derived variable

Overall Reading Summary

An overall reading result was derived by taking the results of the reading task, and adding level 3 / 4+ from the reading comprehension test. This does not make any allowance for differing level 2 grades in the reading task and the reading comprehension test, but does at least provide a reading variable that has a score for all children that were assessed.

sat126 Key stage 1: Reading (overall summary)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 W	496	3.1	3.2	3.2
	1 1	2444	15.5	15.6	18.7
	2 2C	2668	16.9	17.0	35.7
	3 2B	3186	20.2	20.3	56.0
	4 2A	2472	15.6	15.7	71.8
	5 3 / 4+	4433	28.1	28.2	100.0
	Total	15699	99.4	100.0	
Missing	-5 Absent (A)	36	.2		
	-4 Disapplied (D)	61	.4		
	Total	97	.6		
Total		15796	100.0		

Writing Task and Spelling Test

In a similar scenario to the reading assessment, the rubric for the spelling test depends on the result of the writing task. However, the writing task and spelling test yield two completely separate results. The instructions for administering them are as follows:

- The writing task awards levels 1–3, with grades C–A at level 2.
- Those children who achieve level 2 or 3 in the writing task, or whose teacher assessment in writing is level 2 or above must take the spelling test. (N.B. We do not have the results of the teacher assessments for 1997 / 1998 and 1998 / 1999.)
- Children who do not achieve level 2 or 3 in the writing task or level 2 or above in the teacher assessment in writing may also take the spelling test, at the teacher's discretion.
- The spelling test awards levels 1–3, with no grading of level 2. Level 4+ writing can be obtained by taking Key Stage 2 tests.

The instructions for recording the results of the spelling test are as follows:

- Children not entered for the spelling test are shown as X for spelling.
- Children required to take the spelling test who fail to reach level 2 are marked as L for spelling.
- The mark scheme for the spelling test allows for the award of level 1, but since entry at this level is optional, the result should be recorded as X on the results form (but level 1 results may be reported to parents).

Due to the dependence of entry for the spelling test on teacher assessments data that are not generally available and on teacher discretion, logical edits cannot be made in the same way that they were for the reading task and reading comprehension test.

sat130 Key stage 1: Writing task

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 W	1004	6.4	6.4	6.4
	1 1	1937	12.3	12.3	18.7
	2 2C	4937	31.3	31.5	50.2
	3 2B	4397	27.8	28.0	78.2
	4 2A	2401	15.2	15.3	93.5
	5 3 / 4+	1020	6.5	6.5	100.0
	Total	15696	99.4	100.0	
Missing	-5 Absent (A)	39	.2		
	-4 Disapplied (D)	61	.4		
	Total	100	.6		
Total		15796	100.0		

sat135 Key stage 1: Spelling test

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid				
3 2	7333	46.4	69.7	69.7
5 3 / 4+	3189	20.2	30.3	100.0
Total	10522	66.6	100.0	
Missing				
-5 Absent (A)	54	.3		
-4 Disapplied (D)	49	.3		
-3 < level 2 (L)	2431	15.4		
-2 Not entered (X)	2740	17.3		
Total	5274	33.4		
Total	15796	100.0		

A cross-tabulation of these two variables is provided as an aid to understanding the rubric:

sat130 Key stage 1: Writing task * sat135 Key stage 1: Spelling test Crosstabulation

Count

		sat135 Key stage 1: Spelling test						Total
		-5 Absent (A)	-4 Disapplied (D)	-3 < level 2 (L)	-2 Not entered (X)	3 2	5 3 / 4+	
sat130	-5 Absent (A)	29	0	1	4	5	0	39
Key stage	-4 Disapplied (D)	0	49	0	12	0	0	61
1: Writing	0 W	0	0	32	972	0	0	1004
task	1 1	0	0	146	1752	39	0	1937
	2 2C	17	0	2022	0	2782	116	4937
	3 2B	3	0	221	0	3359	814	4397
	4 2A	4	0	9	0	1013	1375	2401
	5 3 / 4+	1	0	0	0	135	884	1020
Total		54	49	2431	2740	7333	3189	15796

Mathematics

Mathematics is assessed by means of a task, which awards level 1, and a test, which awards level 2 (with grades A-C) and level 3. Level 4+ can be achieved by means of Key Stage 2 materials. Since there is only one recorded mark, which is not linked to any other results no edits were required.

sat140 Key stage 1: Mathematics

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 W	421	2.7	2.7	2.7
	1 1	1707	10.8	10.9	13.6
	2 2C	3394	21.5	21.6	35.2
	3 2B	3571	22.6	22.8	58.0
	4 2A	3247	20.6	20.7	78.6
	5 3 / 4+	3350	21.2	21.4	100.0
	Total	15690	99.3	100.0	
Missing	-5 Absent (A)	48	.3		
	-4 Disapplied (D)	58	.4		
	Total	106	.7		
Total		15796	100.0		

Derived Variables

Summary score

This is the sum of the scores obtained in reading, writing and mathematics. The overall reading variable (SAT126) is used as a summary of SAT120 and SAT125. The spelling score is omitted completely as not all children take it and it only awards two levels, so is not very discriminatory. Two versions of the score are provided. For most analyses the prorated score, SAT190b, should be used.

The summary score was created by summing SAT126, SAT130 & SAT140, with missing values recoded to 0. The score based on complete cases only is SAT190a. If any of the components were missing then SAT190a was set to -5. Variable SAT190c is the number of missing components. The prorated score is SAT190b. If more than one component was missing then SAT190b was set to -5. Otherwise SAT190b was scaled by a factor of $3 / (3 - \text{SAT190c})$ and rounded to the nearest integer.

sat190a Key stage 1: Summary score (complete cases)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	253	1.6	1.6	1.6
	1	271	1.7	1.7	3.3
	2	369	2.3	2.4	5.7
	3	738	4.7	4.7	10.4
	4	775	4.9	4.9	15.4
	5	888	5.6	5.7	21.0
	6	1168	7.4	7.5	28.5
	7	1326	8.4	8.5	37.0
	8	1412	8.9	9.0	46.0
	9	1422	9.0	9.1	55.0
	10	1274	8.1	8.1	63.2
	11	1298	8.2	8.3	71.5
	12	1170	7.4	7.5	78.9
	13	1311	8.3	8.4	87.3
	14	1223	7.7	7.8	95.1
	15	765	4.8	4.9	100.0
Total		15663	99.2	100.0	
Missing	-5 1+ components missing	133	.8		
Total		15796	100.0		

sat190b Key stage 1: Summary score (prorated)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	256	1.6	1.6	1.6
	1	271	1.7	1.7	3.4
	2	371	2.3	2.4	5.7
	3	746	4.7	4.8	10.5
	4	775	4.9	4.9	15.4
	5	891	5.6	5.7	21.1
	6	1171	7.4	7.5	28.5
	7	1326	8.4	8.4	37.0
	8	1416	9.0	9.0	46.0
	9	1429	9.0	9.1	55.1
	10	1274	8.1	8.1	63.2
	11	1304	8.3	8.3	71.5
	12	1172	7.4	7.5	79.0
	13	1311	8.3	8.3	87.3
	14	1224	7.7	7.8	95.1
	15	766	4.8	4.9	100.0
	Total	15703	99.4	100.0	
Missing	-5 > 1 component missing	93	.6		
Total		15796	100.0		

sat190c Key stage 1: Number of missing summary score components

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	15663	99.2	99.2	99.2
	1	40	.3	.3	99.4
	2	16	.1	.1	99.5
	3	77	.5	.5	100.0
	Total	15796	100.0	100.0	

Age at testing

The child's age at testing was calculated in both weeks and months. Dates of birth were taken from central ALSPAC records. There is no specific date of testing in the source data (it was most likely not recorded). According to the test booklet, the tasks should have been completed between 1st January and four weeks before the end of the summer term. The tests should have been completed during May. Since the results are supposed to reflect the level the children have reached at the end of the key stage 1 period, the date of testing will be assumed to be 31st July in the relevant years as this will reasonably approximate the end of the summer term.

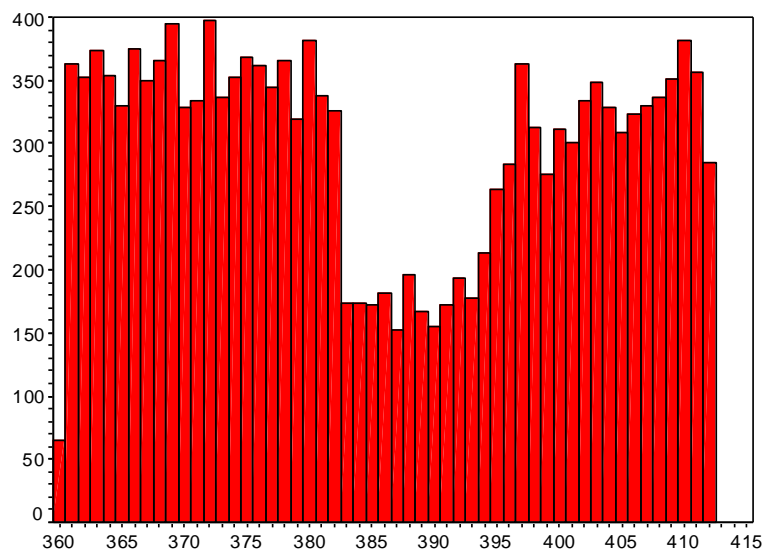
sat195a Key stage 1: Age of child at testing (months)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 78	1	.0	.0	.0
79	1	.0	.0	.0
81	1	.0	.0	.0
82	66	.4	.4	.4
83	1532	9.7	9.7	10.1
84	1611	10.2	10.2	20.3
85	1496	9.5	9.5	29.8
86	1580	10.0	10.0	39.8
87	1521	9.6	9.6	49.4
88	744	4.7	4.7	54.1
89	750	4.7	4.7	58.9
90	850	5.4	5.4	64.3
91	1360	8.6	8.6	72.9
92	1401	8.9	8.9	81.8
93	1428	9.0	9.0	90.8
94	1419	9.0	9.0	99.8
95	6	.0	.0	99.8
96	10	.1	.1	99.9
97	9	.1	.1	99.9
98	1	.0	.0	99.9
99	4	.0	.0	100.0
100	2	.0	.0	100.0
101	2	.0	.0	100.0
106	1	.0	.0	100.0
Total	15796	100.0	100.0	

Age in weeks (SAT195b) is presented separately as a histogram for the children in the correct school year (SAT106a = 1) and as a frequency table for those not in the correct school year (SAT106a = 2):

Children in correct school year:

SAT195B



Key stage 1: Age of child at testing (weeks)

Children not in correct school year (i.e. outliers):

sat195b Key stage 1: Age of child at testing (weeks)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	343	1	2.6	2.6	2.6
	345	1	2.6	2.6	5.1
	355	1	2.6	2.6	7.7
	360	1	2.6	2.6	10.3
	413	1	2.6	2.6	12.8
	414	3	7.7	7.7	20.5
	416	1	2.6	2.6	23.1
	417	3	7.7	7.7	30.8
	418	3	7.7	7.7	38.5
	419	1	2.6	2.6	41.0
	420	2	5.1	5.1	46.2
	421	2	5.1	5.1	51.3
	422	3	7.7	7.7	59.0
	423	1	2.6	2.6	61.5
	424	3	7.7	7.7	69.2
	425	2	5.1	5.1	74.4
	427	1	2.6	2.6	76.9
	432	3	7.7	7.7	84.6
	434	1	2.6	2.6	87.2
	435	1	2.6	2.6	89.7
	439	1	2.6	2.6	92.3
	441	1	2.6	2.6	94.9
	442	1	2.6	2.6	97.4
	463	1	2.6	2.6	100.0
Total		39	100.0	100.0	